

# Reporte de Análisis de Dataset

Archivo analizado: merged\_dataset.csv

Columnas y tipos de datos:

ts	float64
uid	object
id.orig_h	object
id.orig_p	int64
id.resp_h	object
id.resp_p	int64
proto	object
service	object
duration	object
orig_bytes	object
resp_bytes	object
conn_state	object
local_orig	object
local_resp	object
missed_bytes	int64
history	object
orig_pkts	int64
orig_ip_bytes	int64
resp_pkts	int64
resp_ip_bytes	int64
tunnel_parents	object
label	object
detailed-label	object

Valores nulos por columna:

ts	0
uid	0
id.orig_h	0
id.orig_p	0
id.resp_h	0
id.resp_p	0
proto	0
service	24993006
duration	15272073
orig_bytes	15272073
resp_bytes	15272073
conn_state	0
local_orig	25011003
local_resp	25011003
missed_bytes	0
history	25116
orig_pkts	0
orig_ip_bytes	0
resp_pkts	0
resp_ip_bytes	0
tunnel_parents	25011003

label 0  
detailed-label 17954112

Valores únicos por columnas relevantes:

id.orig\_h: 21442 únicos  
id.orig\_p: 65536 únicos  
id.resp\_h: 11654579 únicos  
id.resp\_p: 65435 únicos  
proto: 3 únicos  
service: 7 únicos  
duration: 1080937 únicos  
orig\_bytes: 66187 únicos  
resp\_bytes: 936 únicos  
conn\_state: 13 únicos  
missed\_bytes: 23 únicos  
history: 264 únicos  
orig\_pkts: 207 únicos  
orig\_ip\_bytes: 1858 únicos  
resp\_pkts: 129 únicos  
resp\_ip\_bytes: 1889 únicos  
label: 7 únicos  
detailed-label: 7 únicos

Estadísticas descriptivas (numéricas):

missed\_bytes: mean=76.33, min=0.0, max=1908819480.0

orig\_pkts: mean=13.50, min=0.0, max=66027354.0

orig\_ip\_bytes: mean=590.02, min=0.0, max=1914793266.0

resp\_pkts: mean=0.01, min=0.0, max=9613.0

resp\_ip\_bytes: mean=1.75, min=0.0, max=520116.0

Distribución de la columna 'label':

Benign	8780158
Malicious	7055007
Malicious DDoS	5778154
Malicious PartOfAHorizontalPortScan	3386241
Malicious C&C	8685
Malicious Attack	2755
Malicious FileDownload	3

Distribución de la columna 'detailed-label':

-	8780158
PartOfAHorizontalPortScan	7041695
C&C	8238
Attack	5962

HeartBeat	940
Torii	30
FileDownload	26

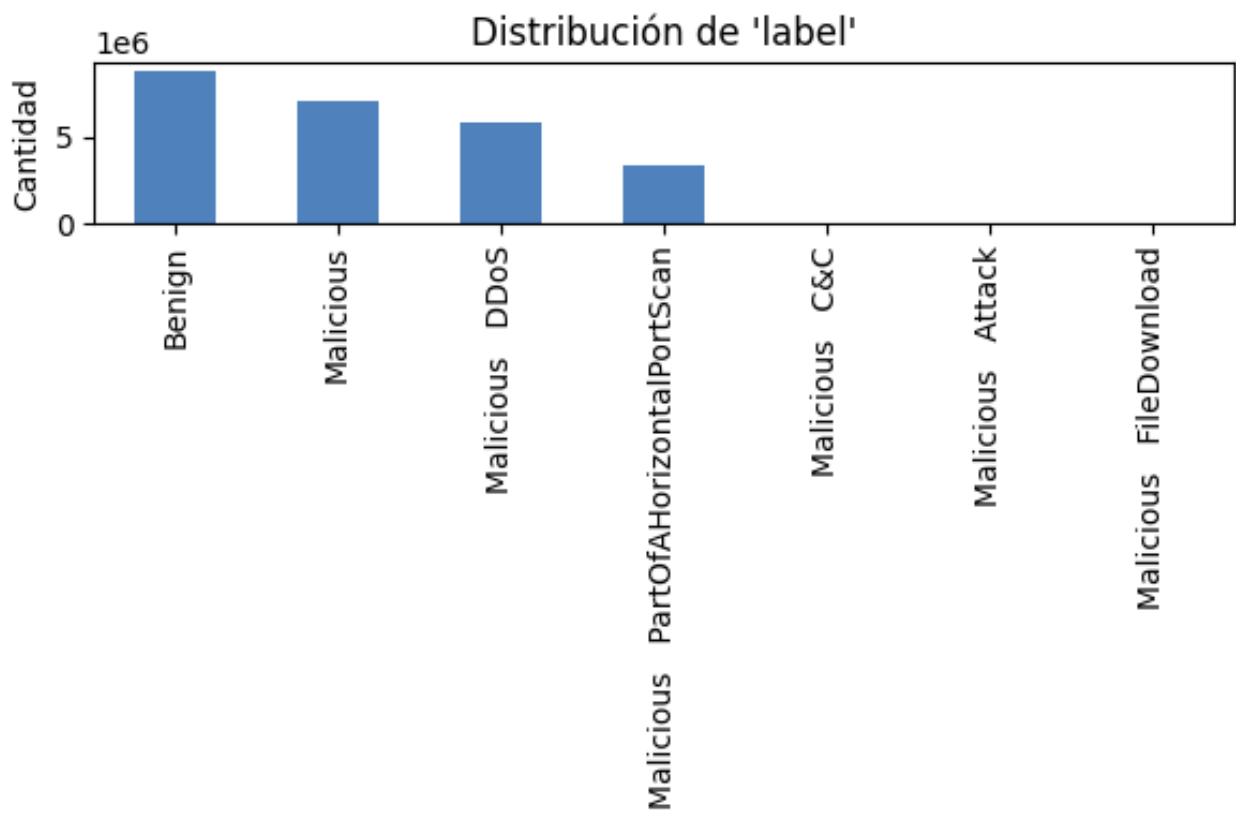
Protocolos más frecuentes (proto):

tcp	24543393
udp	442494
icmp	25116

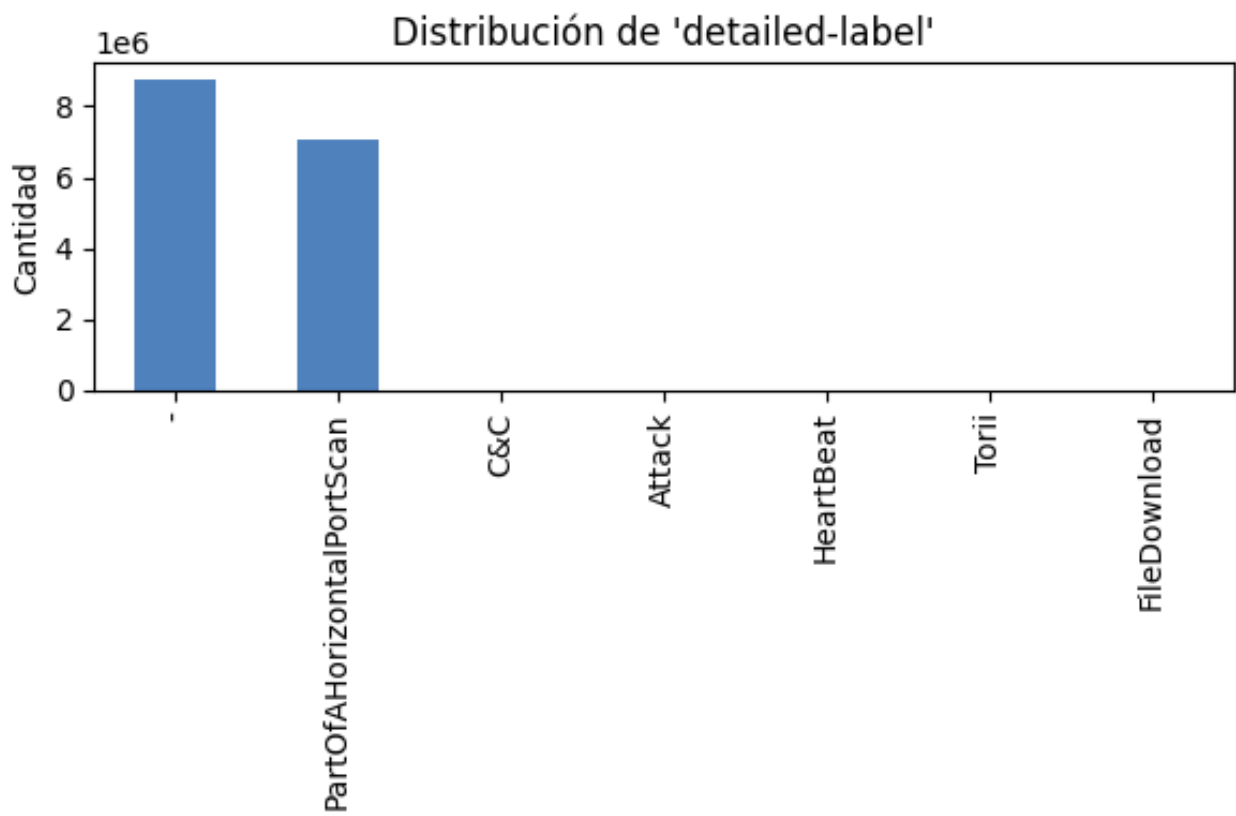
Estados de conexión más frecuentes (conn\_state):

S0	19151121
OTH	3685336
RSTOS0	2119852
SF	33267
REJ	16072
S3	2458
RSTR	1917
RSTO	514
S2	136
S1	131
SH	120
RSTRH	78
SHR	1

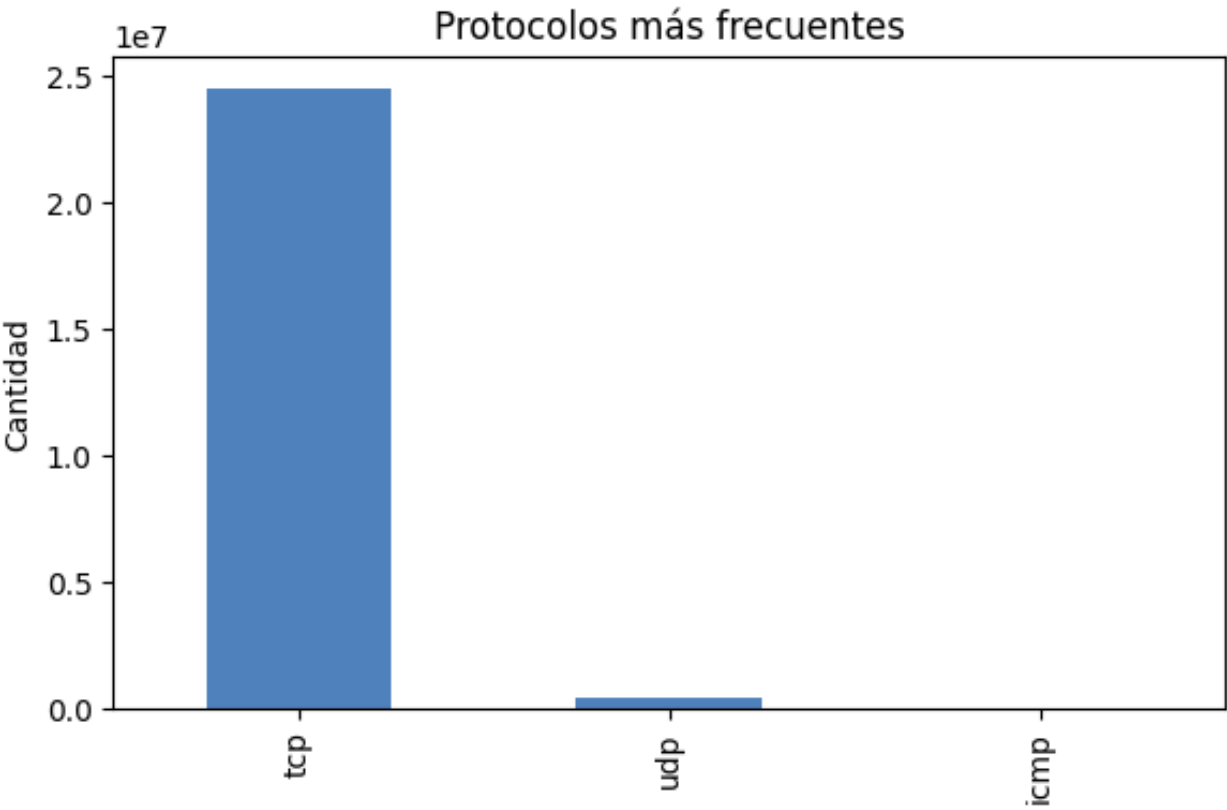
Label dist



Detailed label dist



Proto dist



Conn state dist

