

02 Data Cleaning Report

Canada ↔ Brazil Trade Project

Metric	Value
Final rows	89,412
Final columns	28
Suppressed rows (value = 0)	17
Chapter 98 rows	0
Rows with zero quantity	42,034
UOM errors	2,388
UOM mismatches	0
Value outliers flagged	13,945

Final Dataset Columns

period_m
year
month
quarter
flow
province
country
description
hs2
hs4
hs6
hs8
hs10
section
section_name
chapter_name
value
quantity
uom_text
uom_code
uom_status
has_quantity
is_ch98
is_suppressed
is_outlier
median_value_hs8
lower_bound_hs8
upper_bound_hs8

Sample of Cleaned Dataset

	hs2	hs4	hs6	hs8	hs10	section	section_name	chapter_name	value	quantity	uom_text	uom_code
ubber, o/t hard rubber	00	0000	000000	00000040	0000004008	nan	nan	nan	173	10	Weight in kilograms	KGM
eding 1875 kW	00	0000	000000	00000085	0000008501	nan	nan	nan	2558388	2	Number	NMB
	00	0000	000000	00000084	0000008467	nan	nan	nan	46	0	Blank	BLANK
hreaded, of iron/steel	00	0000	000000	00000073	0000007318	nan	nan	nan	8	0	Weight in kilograms	KGM
out <=3730 W	00	0000	000000	00000085	0000008501	nan	nan	nan	1209	1	Number	NMB
n, within access commitment, nes	00	0000	000000	00000002	0000000202	nan	nan	nan	308144	49957	Weight in kilograms	KGM
	00	0000	000000	00000025	0000002507	nan	nan	nan	1232677	8077863	Weight in kilograms	KGM
en,nes,within access commitment	00	0000	000000	00000002	0000000207	nan	nan	nan	202186	50921	Weight in kilograms	KGM
offal of Nos 0303.91-0303.99	00	0000	000000	00000003	0000000303	nan	nan	nan	371569	21400	Weight in kilograms	KGM
	00	0000	000000	00000023	0000002309	nan	nan	nan	483937	13304	Weight in kilograms	KGM
ug/sweet/spirit,nes,not pufrs	00	0000	000000	00000020	0000002008	nan	nan	nan	207619	131150	Weight in kilograms	KGM
	00	0000	000000	00000017	0000001701	nan	nan	nan	69141	50	Weight in metric tonne	TNE
oly,nes,dspr in a non-aqueous medium	00	0000	000000	00000032	0000003208	nan	nan	nan	99131	1070	Volume in litres	LTR
es	00	0000	000000	00000039	0000003926	nan	nan	nan	5976	0	Blank	BLANK
or,f automotive goods of Ch 87	00	0000	000000	00000040	0000004016	nan	nan	nan	187	0	Blank	BLANK