Уникальные экземпляры столбца «размеры»:

Номер строки	Сам экземпляр	Фотография/ Примечание
0	10	
13	6	
21343	2.5	
53908	100	
24361	10 1500	Две цифры через пробел
24500	0.65 1500	Цифры могут быть дробными
1213	20K1	1213 1211 20К1 1214 1212 1661 1215 1213 30K2 1216 1214 30Ш1 1217 1215 30Ш2 1218 1216 1652 1219 1217 1661 1220 1218 1861 1221 1219 20Ш1 1222 1220 30Ш2 1223 1221 20W2 ЧИСЛО + БУКВЫ [один из: К/Б/Ш/БС] + ЦИФРА Еще встречаются буквы: М (монорельс), ДК (двутавр колонный), ДБ (двутавр
1663	IPE120	балочный). 1657 1655 3561 1658 1656 35Ш2 1659 1657 35Ш1 1660 1658 35K2 1661 1659 35K1 1662 1660 35K1 1663 1661 IPE120 1664 1662 25K1 1665 1663 2551 1666 1664 1462 НУ тут слово IPE и потом идет число
1743	HE200A	1743 1741 HE180A 1744 1742 HE200A 1745 1743 HE200A 1746 1744 HE200A 1747 1745 HE200A 1748 1746 HE200A 1749 1747 HE200A 1750 1748 HE200A 1751 1749 HE200A 1752 1750 HE200A
1761	HE650AA	1757 1755 40Б2 1758 1756 40Б1 1759 1757 45Б1 1760 1758 40Ш2 1761 1759 НЕБ5ОАА 1762 НЕЗООА 1763 1761 НЕ24ОА 1764 1762 НЕ24ОА 1765 1763 НЕ24ОА 1766 1764 НЕ24ОА 1767 1765 40Ш1 В конце идет двойная АА, не понятно опечатка ли это или нет

3061	4x120	3061 3059 4x120 3062 3060 4x100 3063 3061 3.5x90 3064 3062 3x80 3065 3063 3x70 3066 3064 2.5x60 3067 3065 2.5x50 3068 3066 18x32 3070 3068 25 3071 3069 32 Длина х Ширину (видимо)
9895	400x4x32	9891 9889 20 9892 9890 30 9893 9891 32 9894 9892 34 9895 9893 400x4x32 9896 9894 180x2x22 9897 9895 300x3x32 9898 9896 230x2.5x22 9899 9897 355x3.5x25.4 9900 9898 8 9901 9899 16
9924	0.5 9-1000	9924 9922 0.5 9-1000 9925 9923 0.6 9-1000 9926 9924 0.6 9-1000 9927 9925 0.6 9-1000 9928 9926 1.5 9-1000 9929 9927 1.5 9-1000 9930 9928 1.2 9-1000 9931 9929 0.4 9-1000 9932 9930 0.6 9-1000 9933 9931 0.4 9-1000 9933 9931 0.4 9-1000
10195	4 1200x3000 1200x3000	10196 10194 4 1200х3000 1200х3000 10197 10195 4 1200х3000 1200х3000 10198 10196 3 1200х3000 1200х3000 10199 10197 2 1200х3000 1200х3000 10200х 10200 10198 2 1200х3000 1200х3000 10200 10200 1.5 1200х3000 1200х3000 10200 10202 10200 1.5 1200х3000 1200х3000 10202 10200 1.5 1200х3000 1200х3000 10202 10203 10201 1 1200х3000 1200х3000 10202 10203 10201 1 1200х3000 1200х3000 10203 10204 10202 1 1200х3000 1200х3000 10203 10204 10204 0.5 1200x3000 1200x3000 10200x3000 10200x30
10803	508 1200x2000-3500	10802 10800 510 1500x2800 10803 10801 508 1200x2500 10804 606 1000x3500 10805 10803 606 1200x3500 10806 10804 606 1500x3500 10807 10805 606 1500x3500 10808 10806 608 1000x2500 10809 10807 608 1200x2500 10810 10808 606 1000x3500 10811 10809 510 1000x2000 10812 10810 700 1500x4000 Число число х диапазон Какой-то

10935	410	10931 10929 506 1200x3300 10932 10930 204 1000x2600 10933 10931 406 1200x3000 10934 10932 306 1500x3000 10935 10933 410 10936 10934 510 10937 10935 508 1000x2010 10938 10936 510 1200x1500-3500 10939 10937 510 1000x1500-3000 10940 10938 406 1000x2510 Приэтом в каки-то моменты идет просто число без число х число
14936	50 1500-1600x5500-6000	14929 14927 36 1500x6000 14930 14928 8 1500x6000 14931 14929 6 2000x6000 14932 14930 16 1500x6000 14933 14931 16 1500x6000 14934 14932 16 1500x6000 14935 14933 32 1500x6000 14936 14934 20 1500x6000 14937 14935 16 2000x6000 14938 14936 50 1500-1600x5500-6000 14939 14937 30 1500x5500-6000 ЧИСЛО ДИАПАЗОН X ДИАПАЗОН
19768	33x11/30x2	Появляется какая-то дробь между, не понимаю как это интерперетировать
21436	ΠΗ 28x27x0.5	21436 21434 ПН 28x27x0.5 21437 21435 ПП 60x27x0.45 21438 21436 140x100x4 21439 21437 140x140x4 21440 21438 140x140x10 21441 21439 140x140x6 21442 21440 140x140x6 21443 21441 140x140x5 21444 21442 120x80x10 21445 21443 120x80x8 Появляются какие-то буквы ПН и ПП перед число х число х число
22247	C20 0.5x1100(1150)	22247 22245 C20 0.5x1100(1150) 22248 22246 C8 0.45x1150(1200) 22249 22247 C21 0.7x1000(1051) 22250 22248 C21 0.45x1000(1051) 22251 22249 C8 0.45x1150(1200) 22252 22250 C21 0.7x1000(1051) 22253 22251 C20 0.5x1100(1150) 22254 22252 C21 0.7x1000(1051) 22255 22253 C20 0.7x1100(1150) 22256 22254 C21 0.4x1000(1051) 22257 22255 H75 0.7x750 22258 22256 C8 0.7x1150(1200) БУКВЫ[Н или С или НС] ЧИСЛО ЧИСЛОхЧИСЛО м.б. еще какое-то число в скобках
24358	P33	24358 24356 P33 24359 24357 KP70 24360 24358 KP120 24361 24359 10 1500 24362 24360 10 2000 24363 24361 8 2000 БУКВЫ[Р или КР или м.б. К]ЧИСЛО
26210	250 0.5/0.5 1000	26209 26207 100 0.5/0.5 1000 26210 26208 250 0.5/0.5 1000 26211 26209 250 0.5/0.5 1190 26212 26210 240 0.5/0.5 1190 26213 26211 160 0.5/0.5 1190 Между числами через пробел может появиться /

51371	10У	51371 51374 109 51372 51375 59 51373 51376 51 51374 51377 81 51375 51378 81 51376 51376 101 ЧИСЛО БУКВА[П или У]
51503	UPN120	51503 51506 UPN120 51504 51507 149 51505 51508 149 51506 51509 149 51507 51510 149 51508 51511 14П UPN ЧИСЛО
53173	СШК 22-600	S3171 S3174 Z0X1000
53266	СШД 190-750/927	53265 53268 СШД 190-750/927
53409	ПШСТ-1020х14-КО	53309 53312 ПШСТ-1020x14-КО 53310 53313 ПШСТ-1020x11-КО 53311 53314 ПШСТ-1020x12-КО ПШСТ-ЧИСЛОхЧИСЛО-КО
53241	Л5-УМ	53238 53241 Л5-УМ 53239 53242 Л5-УМ 53240 53243 Л5-УМ Начинается что-то вообще не понятное
53341	КО еще есть КОх2	53341 53344 KO 53342 53345 KOX2 53343 53346 0.35 118
С 53913 и до конца	Шаблоны: Omnia-46 ЦИФРА OK-46 ЦИФРА MP-3 или MP-3С и ЦИФРА E308-L-16 ЧИСЛО х ЧИСЛО ЛЭ3-46 ЦИФРА	53913 53916 Omnia-46 4 53914 53917 OK-46 3 53915 53918 MP-3 5 53916 53919 E308L-16 2x350 53917 53920 E308L-16 2.6x350 53918 53921 J33-46 3 53919 53922 OK-46 4 53920 53923 Omnia-46 3 53921 53924 MP-3C 4 53922 53925 MP-3C 3 53923 53926 MP-3C 3 53924 53927 MP-3C 3

Итоговый список шаблонов для парсера (от самого сложного к простому)

Это рекомендуемый порядок проверки, так как более специфичные шаблоны должны проверяться раньше более общих.

- 1.**Профнастил/Сетки (с буквами и скобками):** C20 0.5x1100(1150)
- **2.Листовой прокат (толщина + размеры):** 5 1500x6000
- 3.**Сложные размеры с диапазоном:** | 508 1200x2000-3500
- 4. Сложные размеры с дробью: 33x11/30x2
- **5.Сложные размеры со слэшем:** 250 0.5/0.5 1000
- 6.Листовой прокат (толщина + диапазон): 0.5 9-1000
- 7.**Профили с префиксом ПН/ПП:** ПН 28x27x0.5
- 8.**Европейские профили:** IPE120, HE200A, UPN120
- 9.**Профили по ГОСТ (балки/швеллеры):** | 20K1, | 16Б2, | 24M, | 40K5
- 10.**Рельсы:** Р33, КР70
- 11.**Электроды/Бренды:** Omnia-46 4, E308L-16 2x350
- 12.Габариты (3 размера): 400х4х32
- 13.**Габариты (2 размера):** 4х120
- 14. Число с суффиксом У/П: 10У, 14П
- 15.**Простое число (целое/дробное):** 10, 6.5