

Overview

Brought to you by [YData](#)

Overview

Alerts 8

Reproduction

Alerts

cap-diameter is highly overall correlated with stem-height and 1 other fields	High correlation
has-ring is highly overall correlated with ring-type	High correlation
ring-type is highly overall correlated with has-ring	High correlation
stem-height is highly overall correlated with cap-diameter and 1 other fields	High correlation
stem-width is highly overall correlated with cap-diameter and 1 other fields	High correlation
does-bruise-or-bleed is highly imbalanced (52.7%)	Imbalance
has-ring is highly imbalanced (54.5%)	Imbalance
ring-type is highly imbalanced (61.0%)	Imbalance

Variables

Select Columns

[class](#)

Categorical

Distinct	2
Distinct (%)	< 0.1%

p
e

1007390
863807

Missing	0
Missing (%)	0.0%
Memory size	16.1 MiB

More details

Overview

Categories

Words

Characters

Common Values

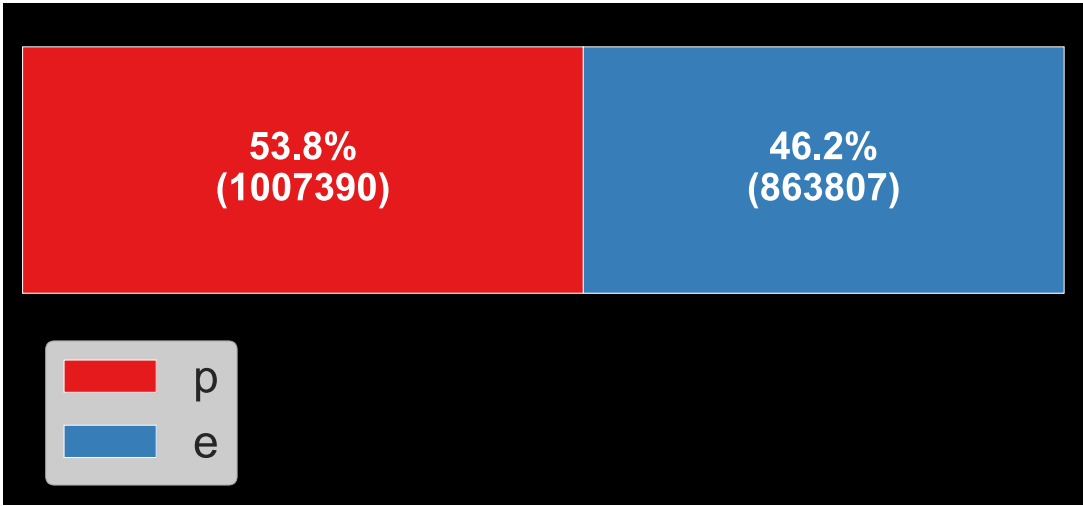
Value	Count	Frequency (%)
p	1007390	53.8%
e	863807	46.2%

Length



Histogram of lengths of the category

Common Values (Plot)

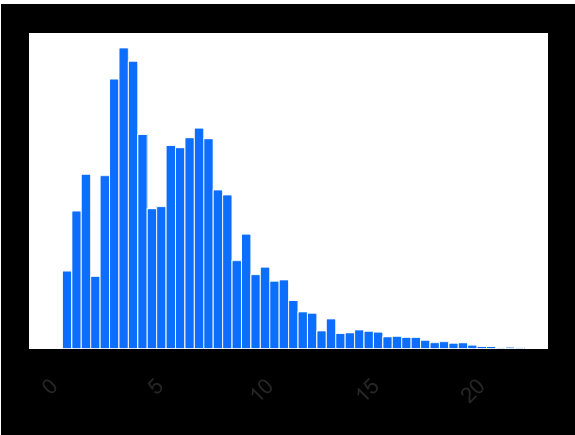


cap-diameter

Real number (\mathbb{R})

High correlation

Distinct	2168	Minimum	0.1
Distinct (%)	0.1%	Maximum	22.38
Missing	0	Zeros	0
Missing (%)	0.0%	Zeros (%)	0.0%
Infinite	0	Negative	0
Infinite (%)	0.0%	Negative (%)	0.0%
Mean	6.1094307	Memory size	28.6 MiB



More details

Statistics

Histogram

Common values

Extreme values

Quantile statistics

Minimum	0.1
5-th percentile	1.42
Q1	3.43

Descriptive statistics

Standard deviation	3.5790745
Coefficient of variation (CV)	0.58582782
Kurtosis	1.4376428

median	5.7	Mean	6.1094307
Q3	7.96	Median Absolute Deviation (MAD)	2.27
95-th percentile	12.95	Skewness	1.0601828
Maximum	22.38	Sum	11431948
Range	22.28	Variance	12.809774
Interquartile range (IQR)	4.53	Monotonicity	Not monotonic

cap-shape

Categorical

Distinct	7
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	16.1 MiB

x	968485
f	430569
s	237811
b	127579
c	55832
Other values ...	50921

More details

Overview

Categories

Words

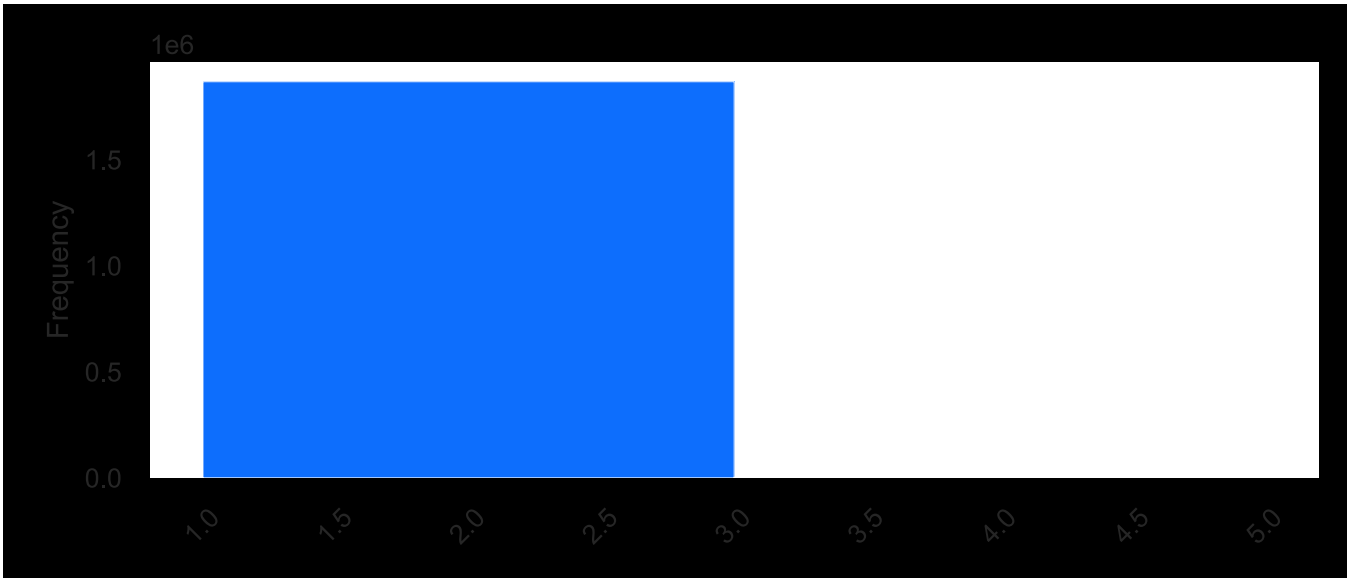
Characters

Common Values

Value	Count	Frequency (%)
x	968485	51.8%
f	430569	23.0%
s	237811	12.7%
b	127579	6.8%

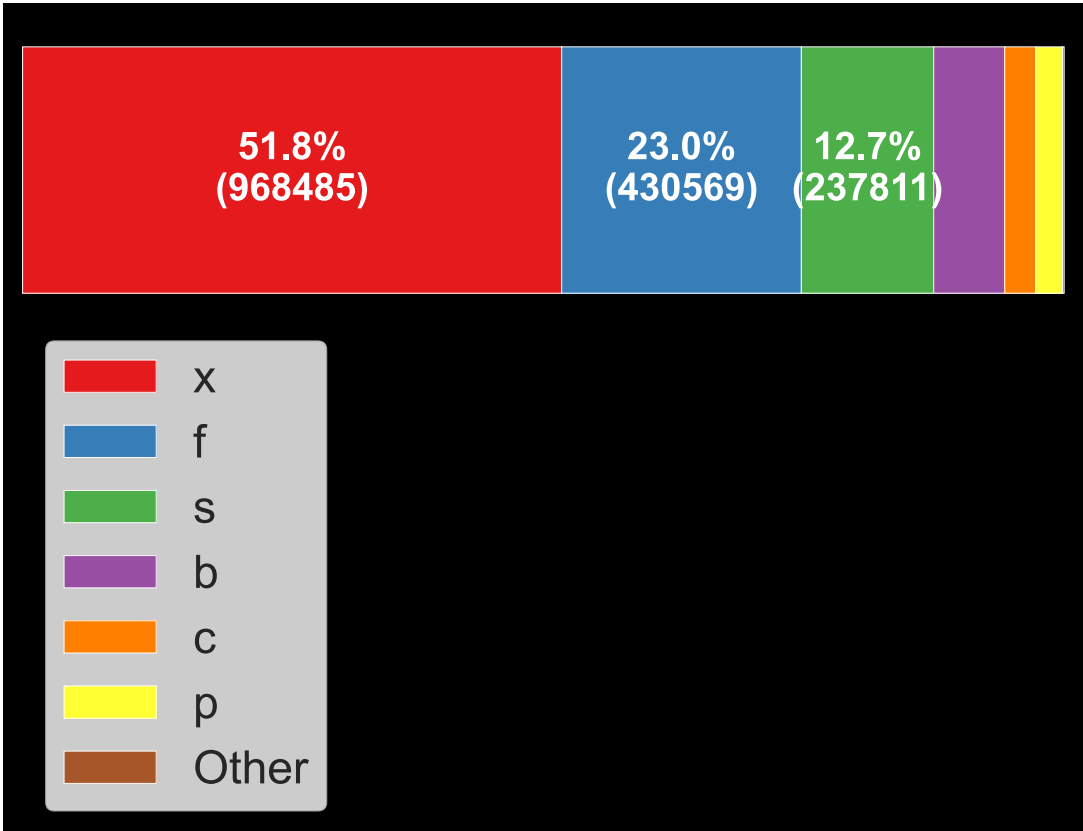
Value	Count	Frequency (%)
c	55832	3.0%
p	47971	2.6%
Other	2950	0.2%

Length



Histogram of lengths of the category

Common Values (Plot)



cap-surface

Categorical

Distinct	11
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	16.1 MiB

t	403882
s	333707
h	209232
g	194088
d	193540
Other values ...	536748

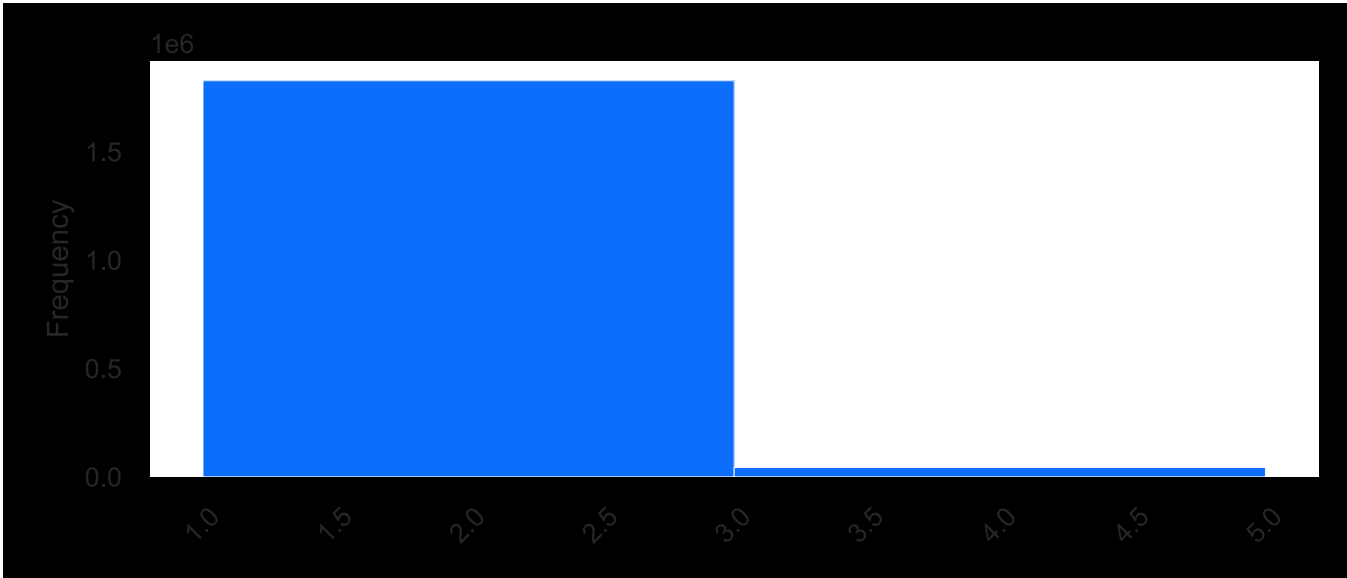
More details

- Overview
- Categories
- Words
- Characters

Common Values

Value	Count	Frequency (%)
t	403882	21.6%
s	333707	17.8%
h	209232	11.2%
g	194088	10.4%
d	193540	10.3%
y	166235	8.9%
e	87418	4.7%
k	84699	4.5%
w	83648	4.5%
i	71413	3.8%

Length



Histogram of lengths of the category

cap-color
Categorical

Distinct	8
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	16.1 MiB

n	856261
w	226251
y	195238
Other	182188
e	126895
Other values ...	284364

More details

Overview

Categories

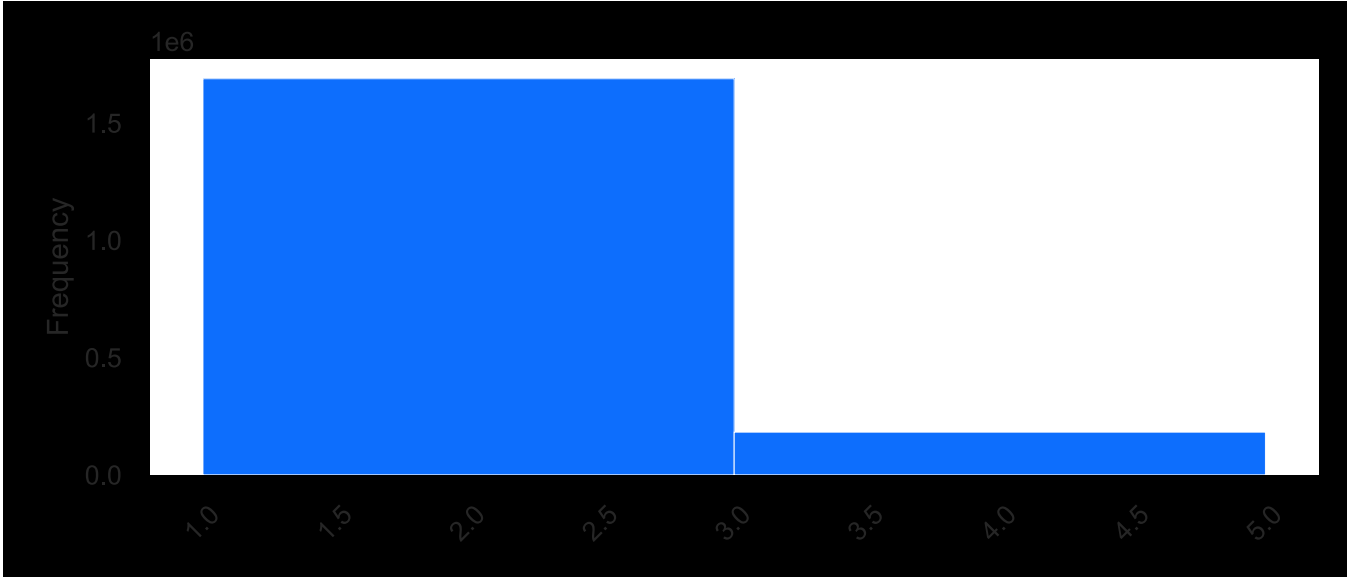
Words

Characters

Common Values

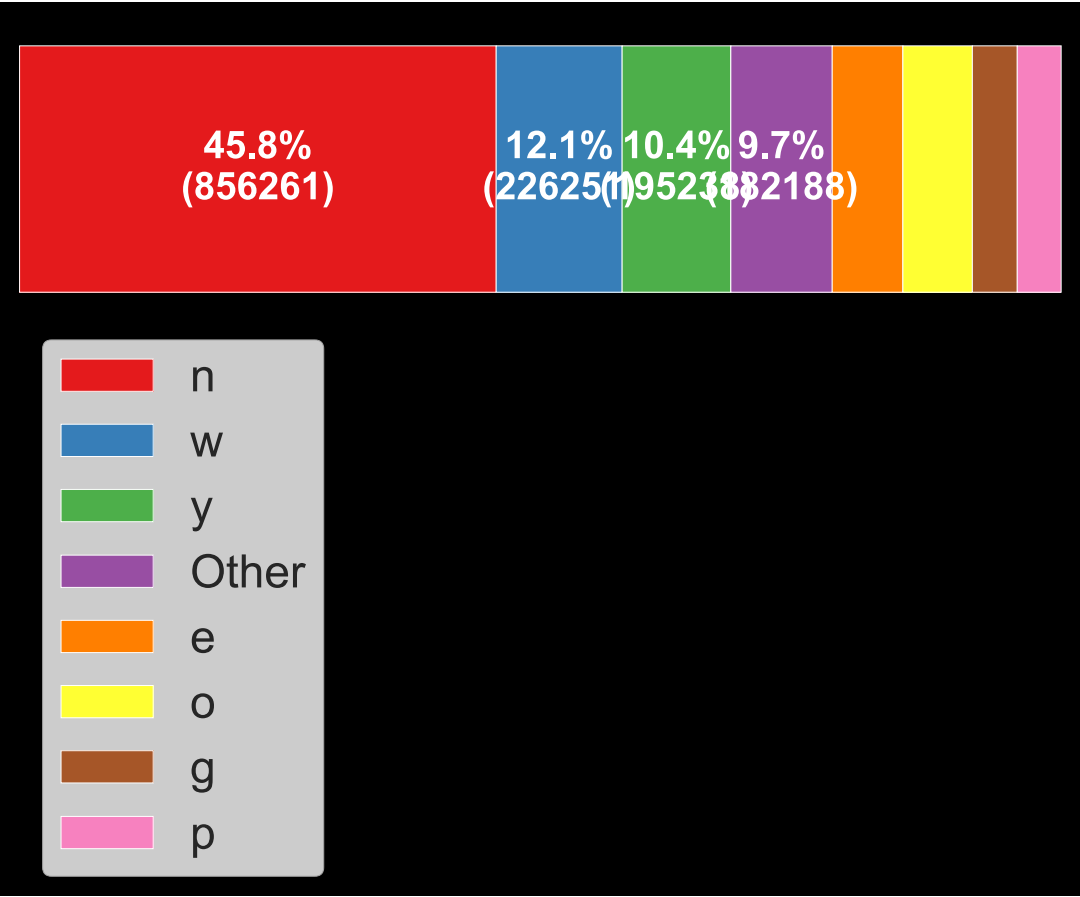
Value	Count	Frequency (%)
n	856261	45.8%
w	226251	12.1%
y	195238	10.4%
Other	182188	9.7%
e	126895	6.8%
o	124856	6.7%
g	80645	4.3%
p	78863	4.2%

Length



Histogram of lengths of the category

Common Values (Plot)



does-bruise-or-bleed

Categorical

Imbalance

Distinct	3
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	16.1 MiB

f	1471350
t	399785
Other	62

More details

Overview

Categories

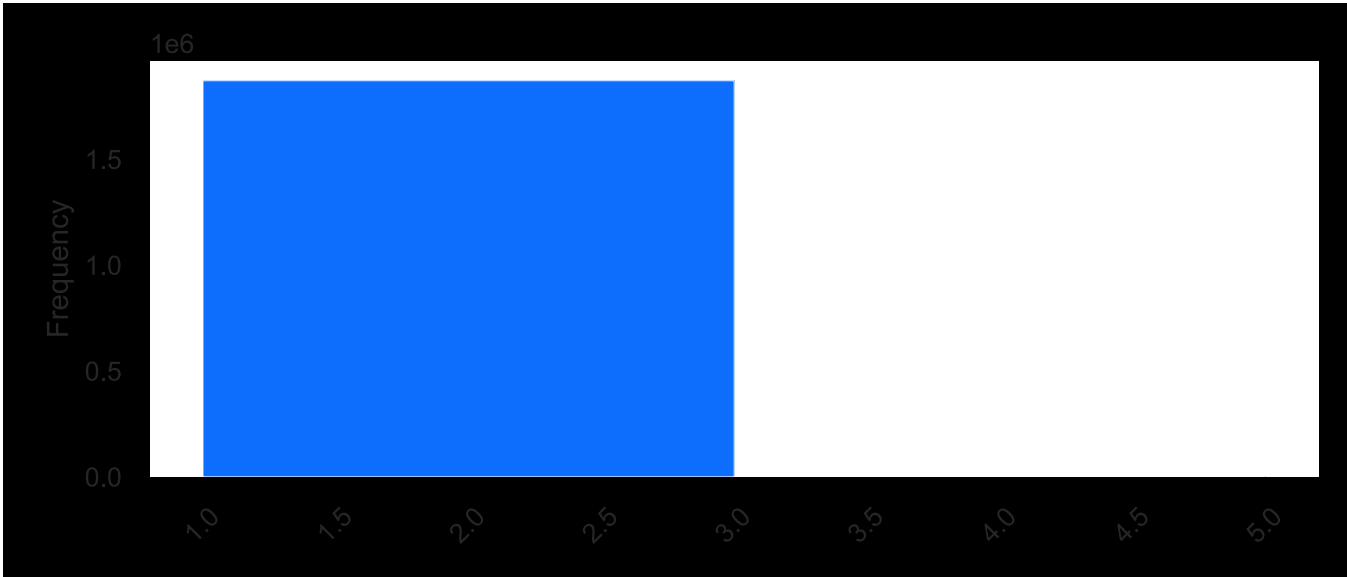
Words

Characters

Common Values

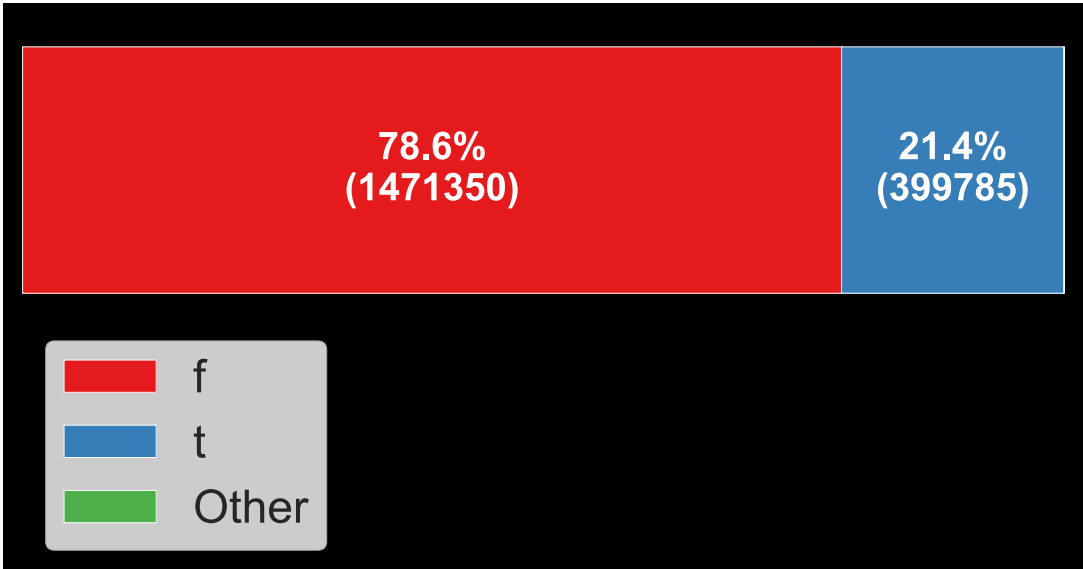
Value	Count	Frequency (%)
f	1471350	78.6%
t	399785	21.4%
Other	62	< 0.1%

Length



Histogram of lengths of the category

Common Values (Plot)



gill-attachment

Categorical

Distinct	7
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	16.1 MiB

a	511198
d	417441
x	254275
s	241689
e	211634
Other values ...	234960

More details

Overview

Categories

Words

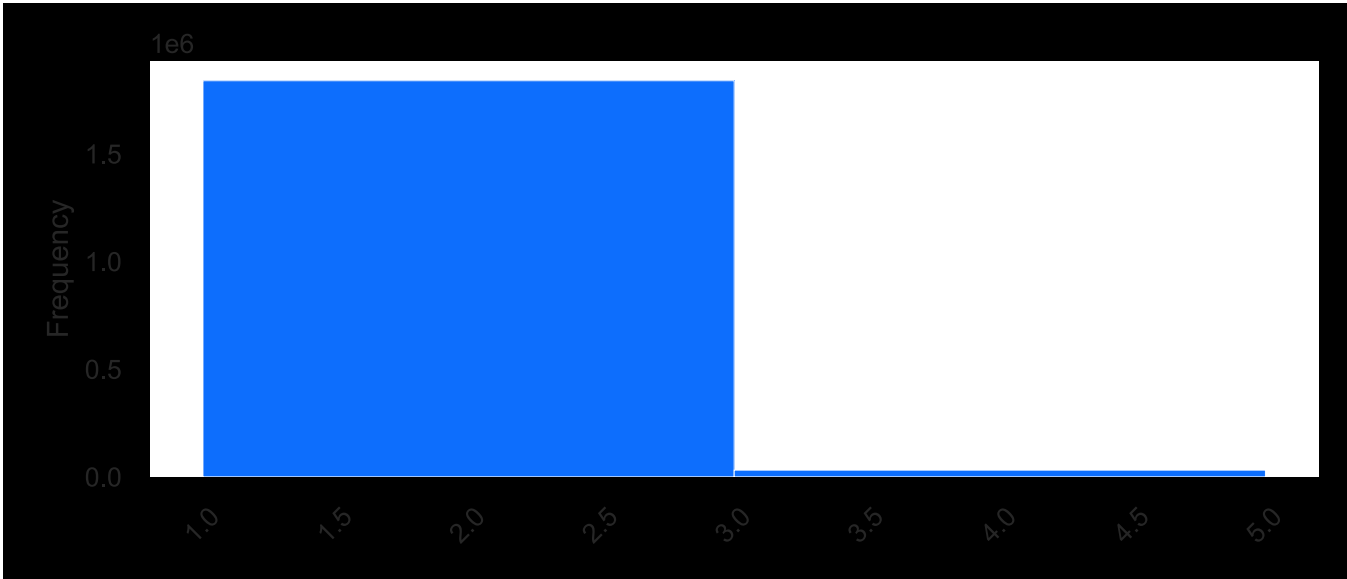
Characters

Common Values

Value	Count	Frequency (%)
a	511198	27.3%
d	417441	22.3%
x	254275	13.6%

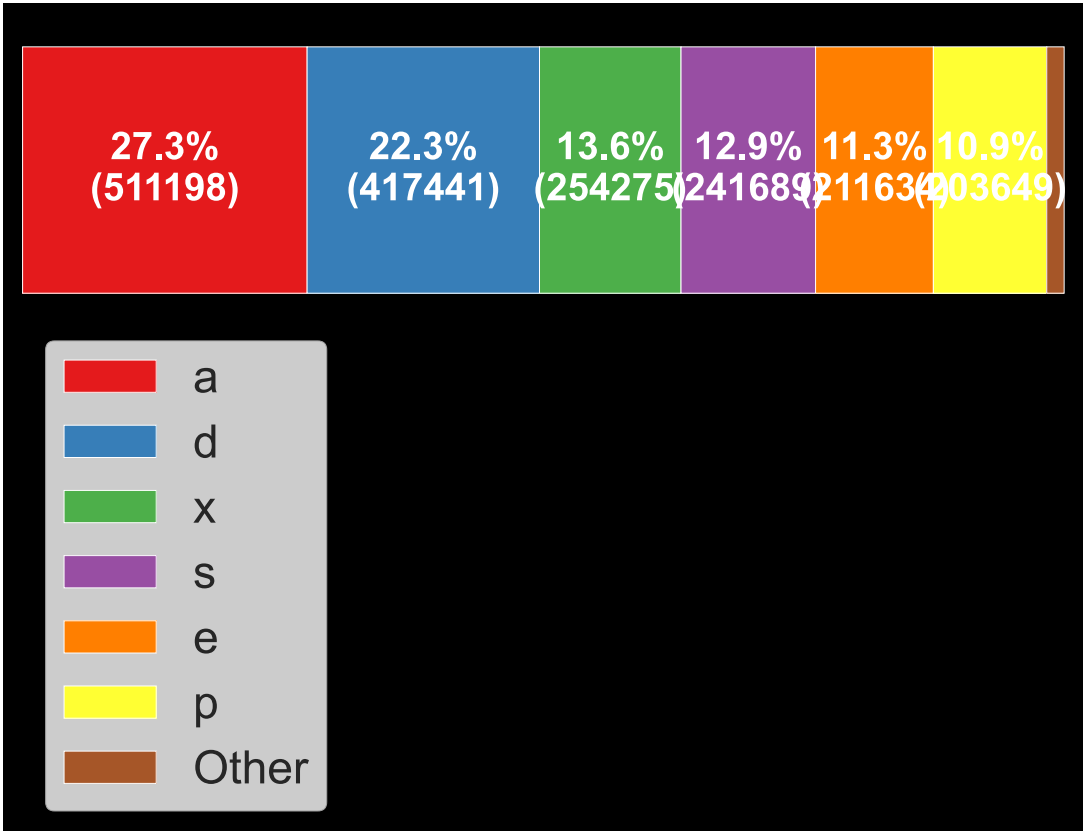
Value	Count	Frequency (%)
s	241689	12.9%
e	211634	11.3%
p	203649	10.9%
Other	31311	1.7%

Length



Histogram of lengths of the category

Common Values (Plot)



gill-color

Categorical

Distinct	7
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	16.1 MiB

w	563136
n	363632
y	314648
p	276287
Other	147526
Other values ...	205968

More details

Overview

Categories

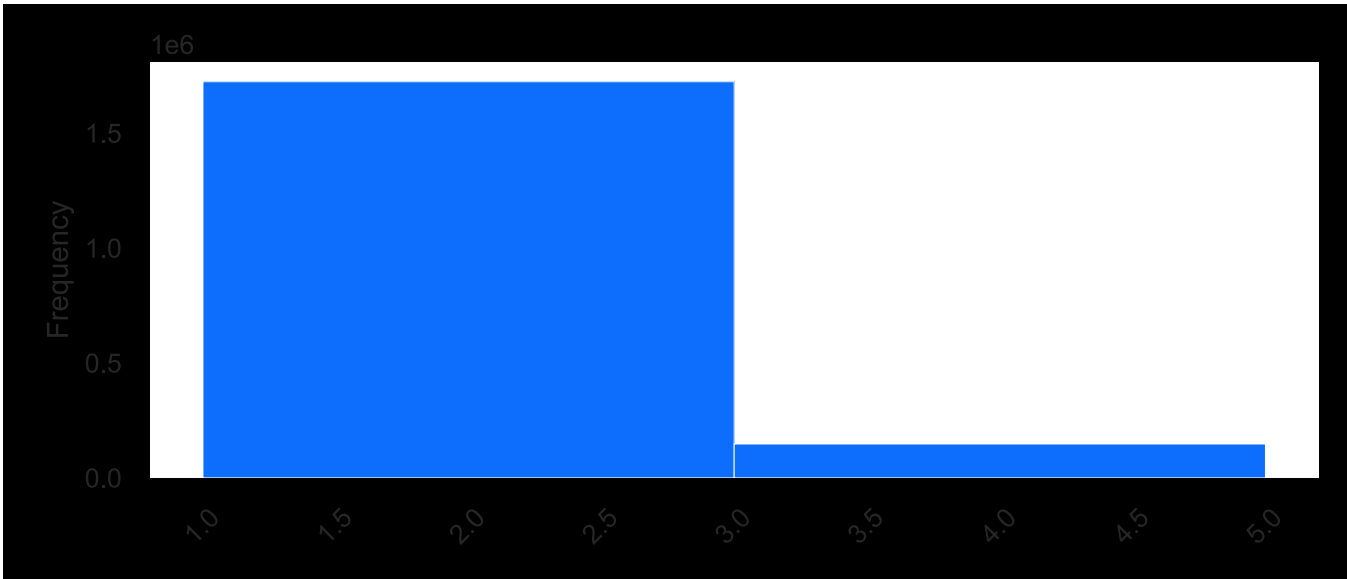
Words

Characters

Common Values

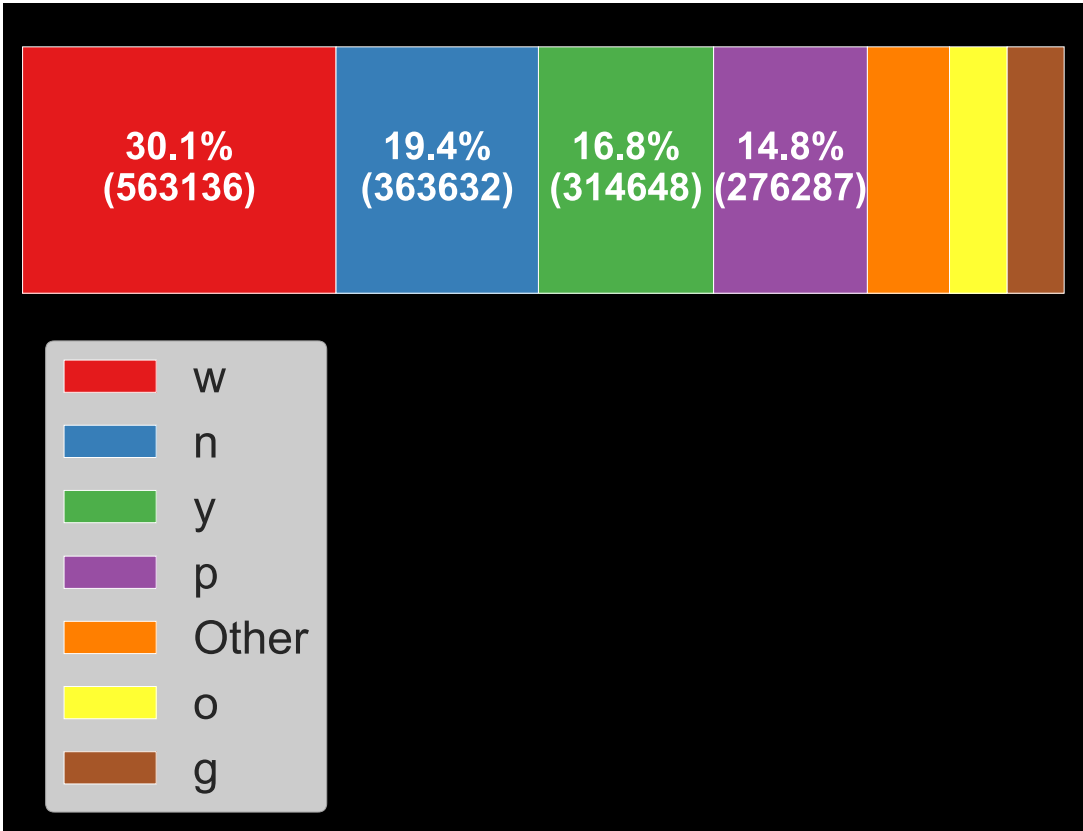
Value	Count	Frequency (%)
w	563136	30.1%
n	363632	19.4%
y	314648	16.8%
p	276287	14.8%
Other	147526	7.9%
o	103595	5.5%
g	102373	5.5%

Length



Histogram of lengths of the category

Common Values (Plot)

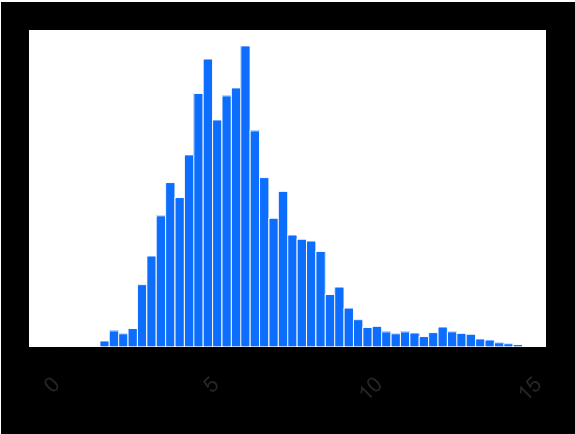


stem-height

Real number (\mathbb{R})

High correlation

Distinct	1361	Minimum	0
Distinct (%)	0.1%	Maximum	14.75
Missing	0	Zeros	222
Missing (%)	0.0%	Zeros (%)	< 0.1%
Infinite	0	Negative	0
Infinite (%)	0.0%	Negative (%)	0.0%
Mean	5.9834298	Memory size	28.6 MiB



More details

Statistics

Histogram

Common values

Extreme values

Quantile statistics

Minimum	0
5-th percentile	3.21
Q1	4.58
median	5.7
Q3	7.05
95-th percentile	9.93
Maximum	14.75
Range	14.75
Interquartile range (IQR)	2.47

Descriptive statistics

Standard deviation	2.0943828
Coefficient of variation (CV)	0.35003048
Kurtosis	1.7179336
Mean	5.9834298
Median Absolute Deviation (MAD)	1.2
Skewness	1.0745814
Sum	11196176
Variance	4.3864393
Monotonicity	Not monotonic

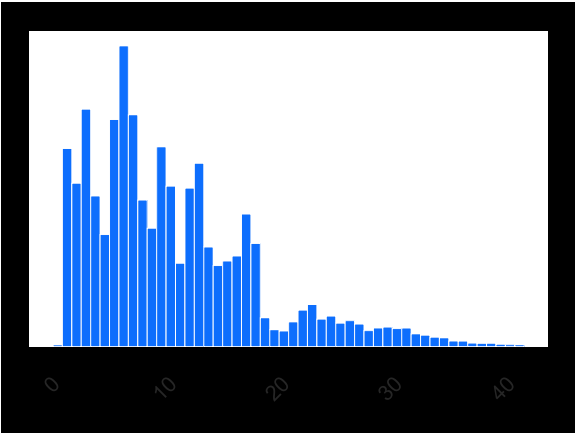
stem-width

Real number (\mathbb{R})

High correlation

Distinct	4024
Distinct (%)	0.2%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Mean	10.651495

Minimum	0
Maximum	41.8
Zeros	207
Zeros (%)	< 0.1%
Negative	0
Negative (%)	0.0%
Memory size	28.6 MiB



More details

Statistics

Histogram

Common values

Extreme values

Quantile statistics

Minimum	0
5-th percentile	1.62
Q1	5.36
median	9.02
Q3	14.38
95-th percentile	26.67
Maximum	41.8
Range	41.8
Interquartile range (IQR)	9.02

Descriptive statistics

Standard deviation	7.5872942
Coefficient of variation (CV)	0.71232197
Kurtosis	1.2673351
Mean	10.651495
Median Absolute Deviation (MAD)	4.31
Skewness	1.1834156
Sum	19931046
Variance	57.567034
Monotonicity	Not monotonic

stem-color

Categorical

Distinct	7
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	16.1 MiB

n	662583
w	611107
y	212984
Other	123297
o	96201
Other values ...	165025

More details

Overview

Categories

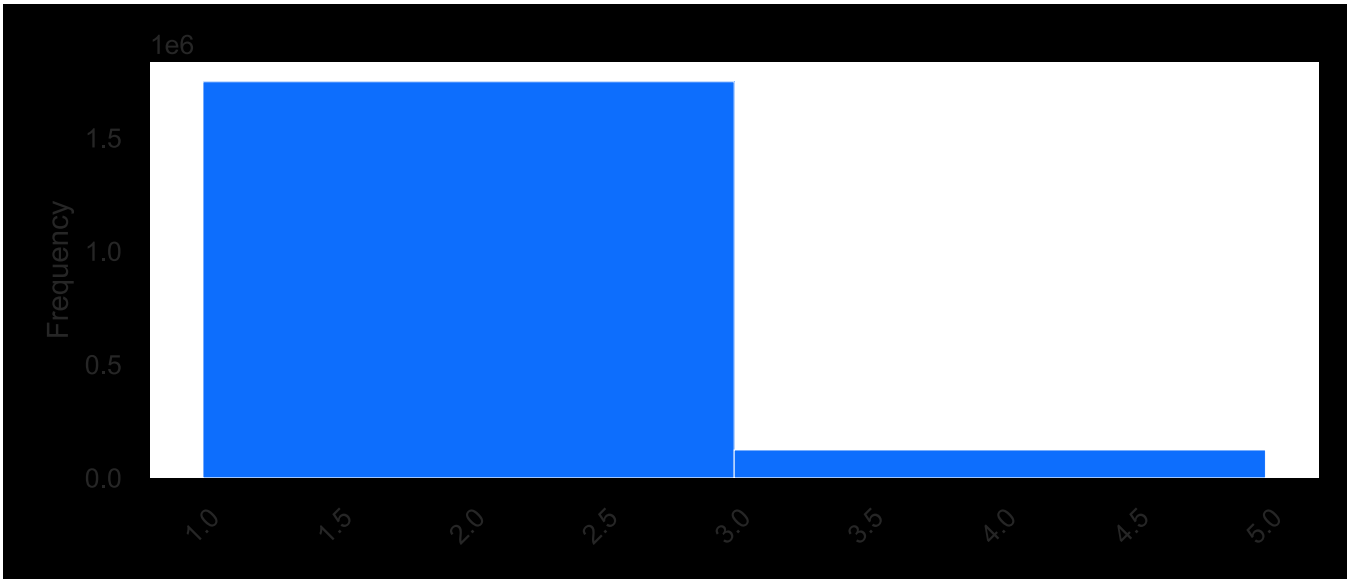
Words

Characters

Common Values

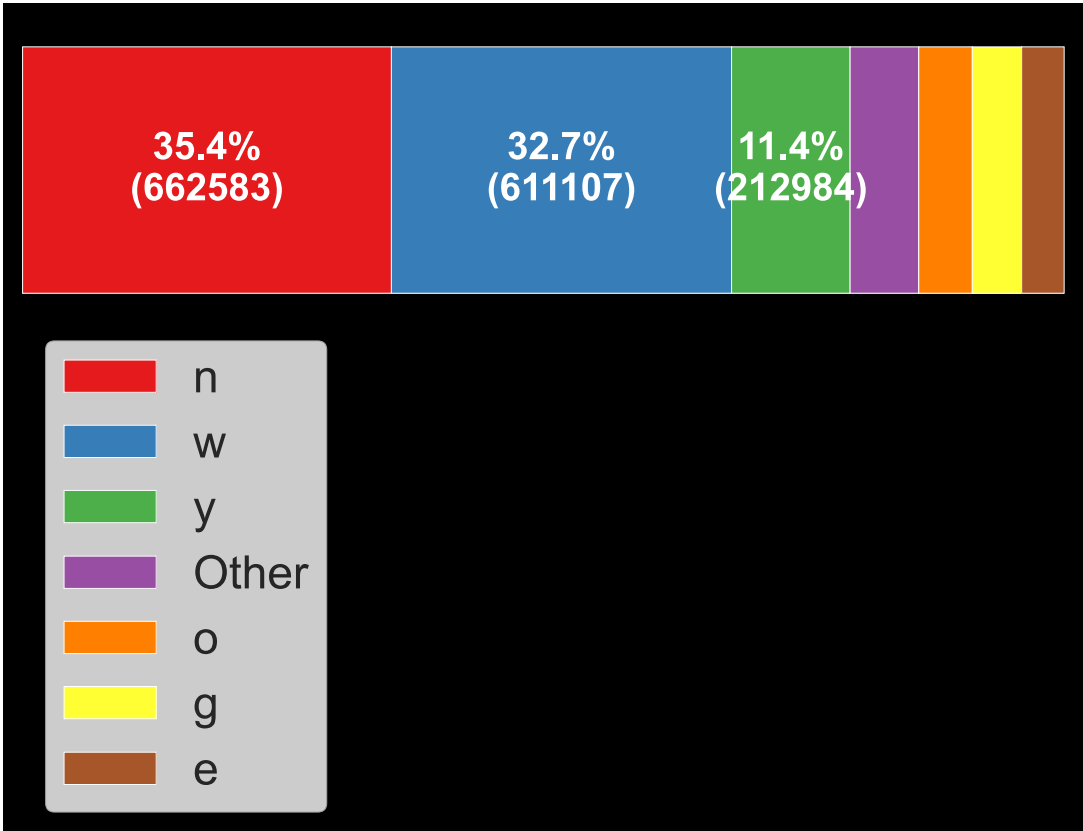
Value	Count	Frequency (%)
n	662583	35.4%
w	611107	32.7%
y	212984	11.4%
Other	123297	6.6%
o	96201	5.1%
g	88660	4.7%
e	76365	4.1%

Length



Histogram of lengths of the category

Common Values (Plot)



has-ring

Categorical

High correlation Imbalance

Distinct	3
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	16.1 MiB

f	1497823
t	373304
Other	70

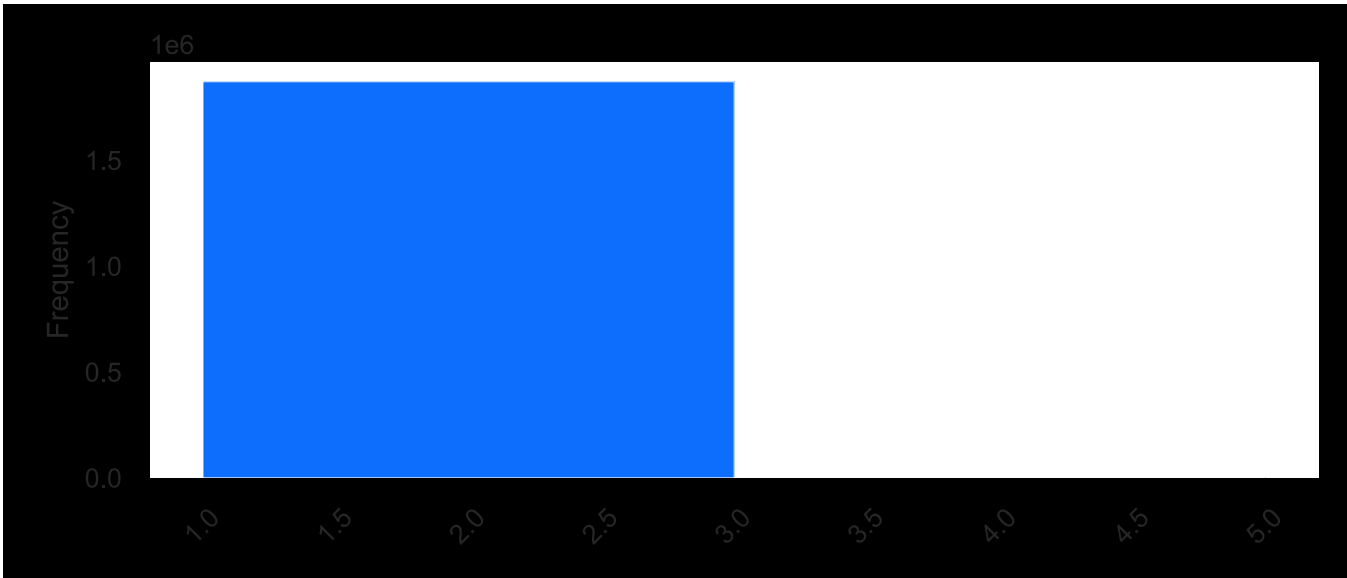
More details

OverviewCategoriesWordsCharacters

Common Values

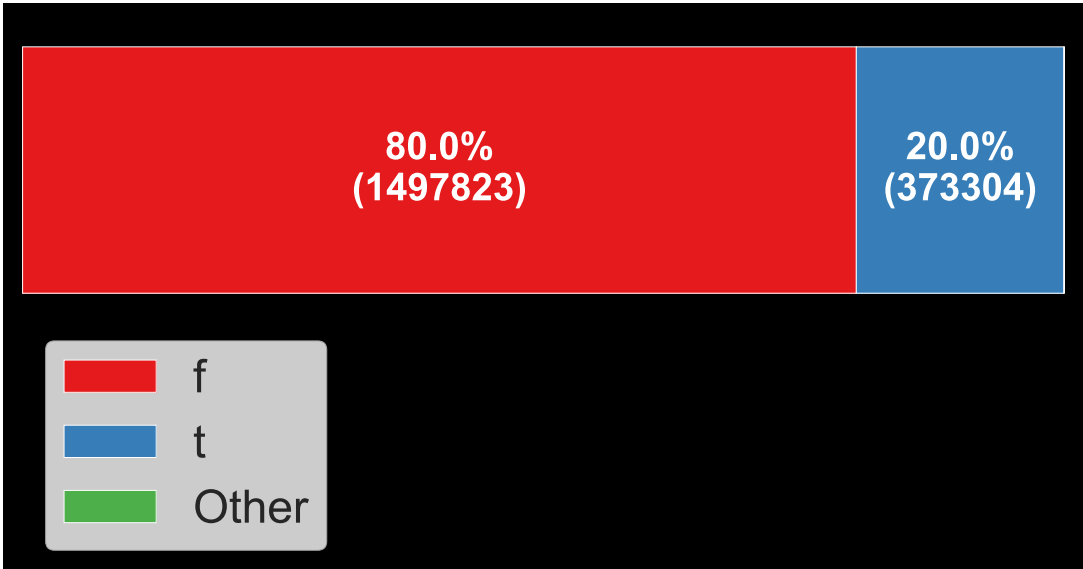
Value	Count	Frequency (%)
f	1497823	80.0%
t	373304	20.0%
Other	70	< 0.1%

Length



Histogram of lengths of the category

Common Values (Plot)



High correlation Imbalance

Distinct	7
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	16.1 MiB

f	1535835
z	80843
e	79894
r	67409
l	55705
Other values ...	51511

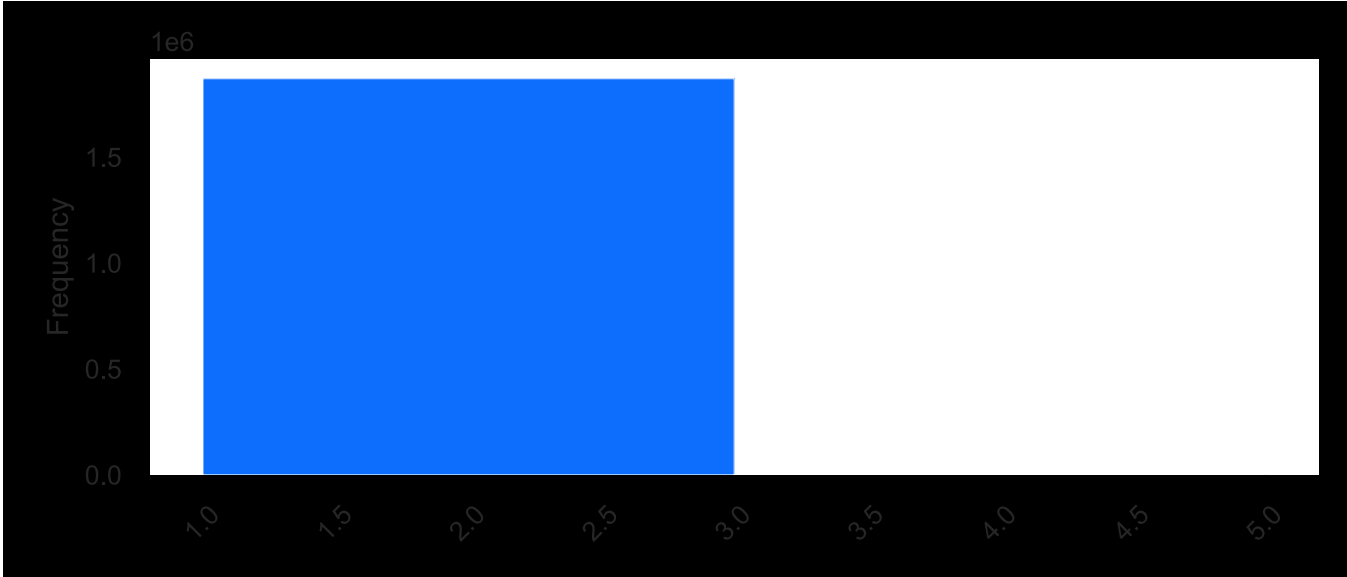
More details

Overview Categories Words Characters

Common Values

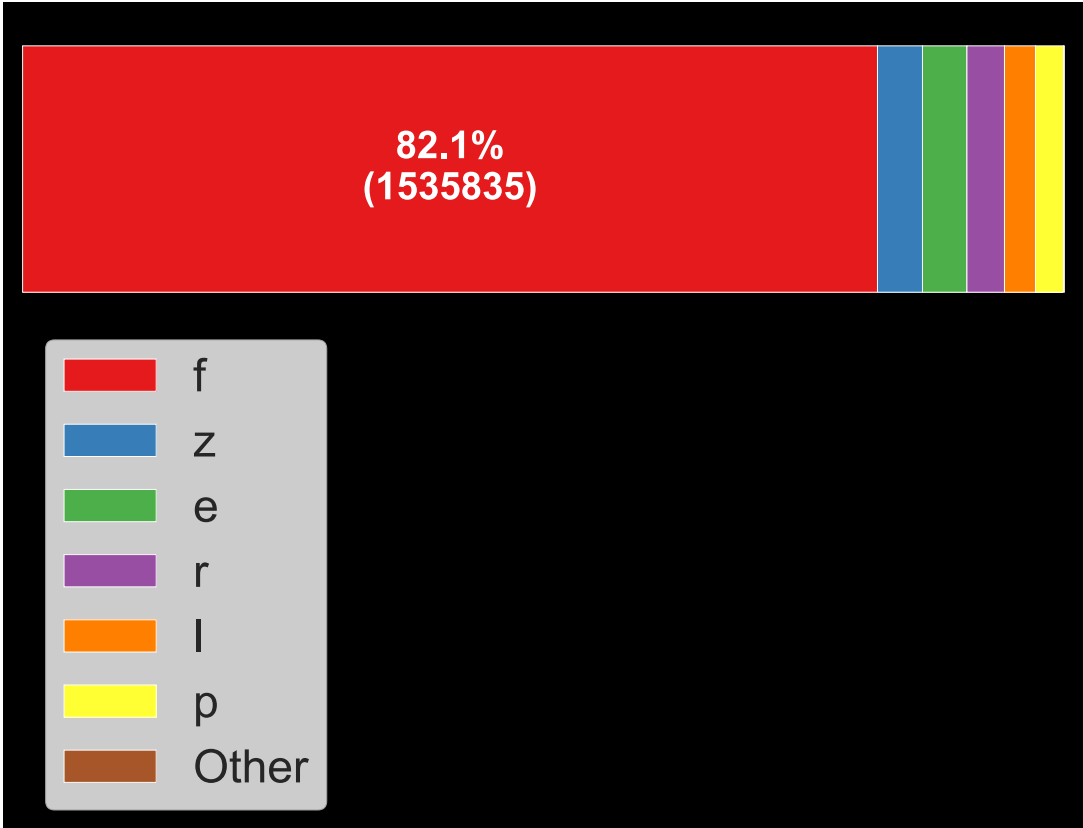
Value	Count	Frequency (%)
f	1535835	82.1%
z	80843	4.3%
e	79894	4.3%
r	67409	3.6%
l	55705	3.0%
p	50668	2.7%
Other	843	< 0.1%

Length



Histogram of lengths of the category

Common Values (Plot)



habitat

Categorical

Distinct	6
----------	---

d	1249771
g	260920

Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	16.1 MiB

m 127341
l 111658
h 101790

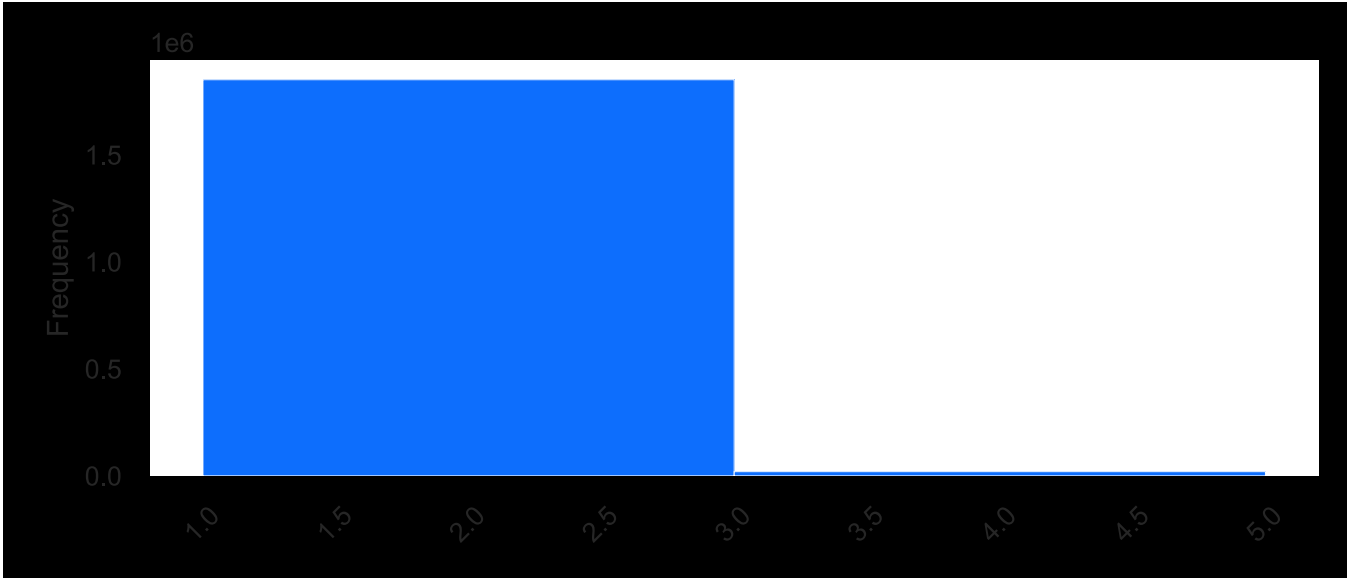
More details

OverviewCategoriesWordsCharacters

Common Values

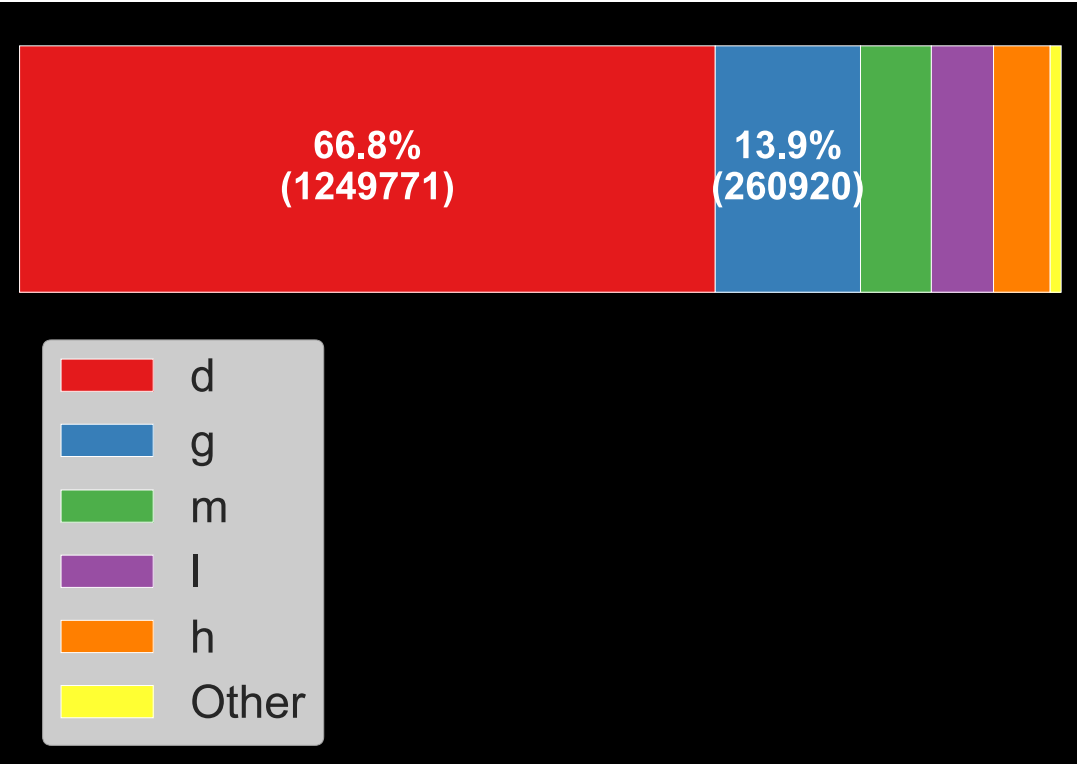
Value	Count	Frequency (%)
d	1249771	66.8%
g	260920	13.9%
m	127341	6.8%
l	111658	6.0%
h	101790	5.4%
Other	19717	1.1%

Length



Histogram of lengths of the category

Common Values (Plot)



season

Categorical

Distinct	4
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	16.1 MiB

a	942037
u	716125
w	160833
s	52202

More details

Overview

Categories

Words

Characters

Common Values

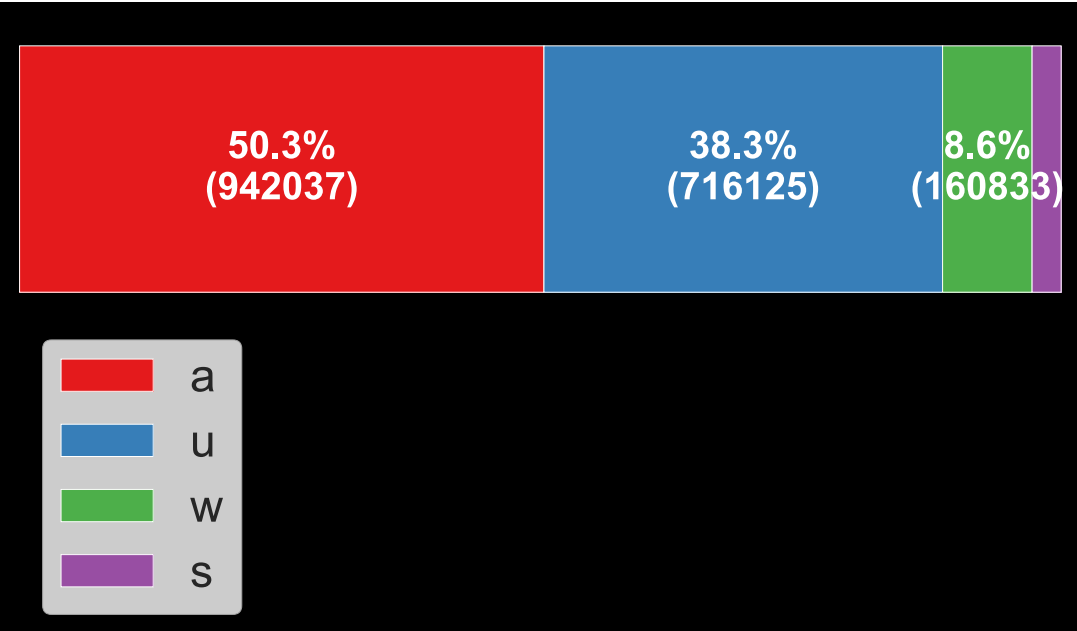
Value	Count	Frequency (%)
a	942037	50.3%
u	716125	38.3%
w	160833	8.6%
s	52202	2.8%

Length

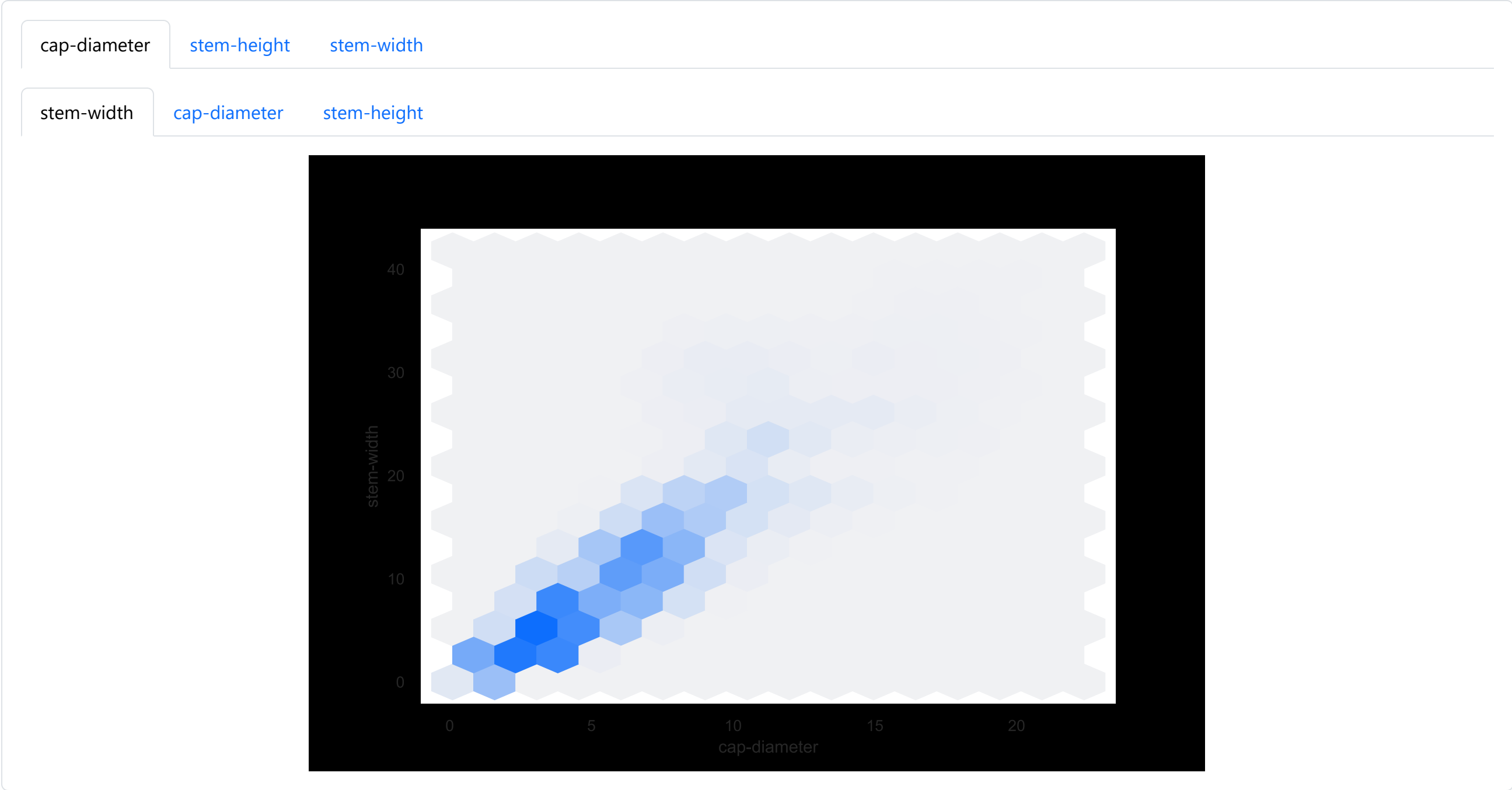


Histogram of lengths of the category

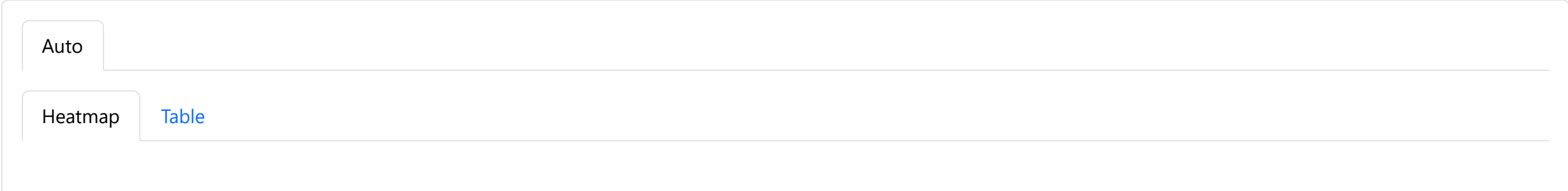
Common Values (Plot)

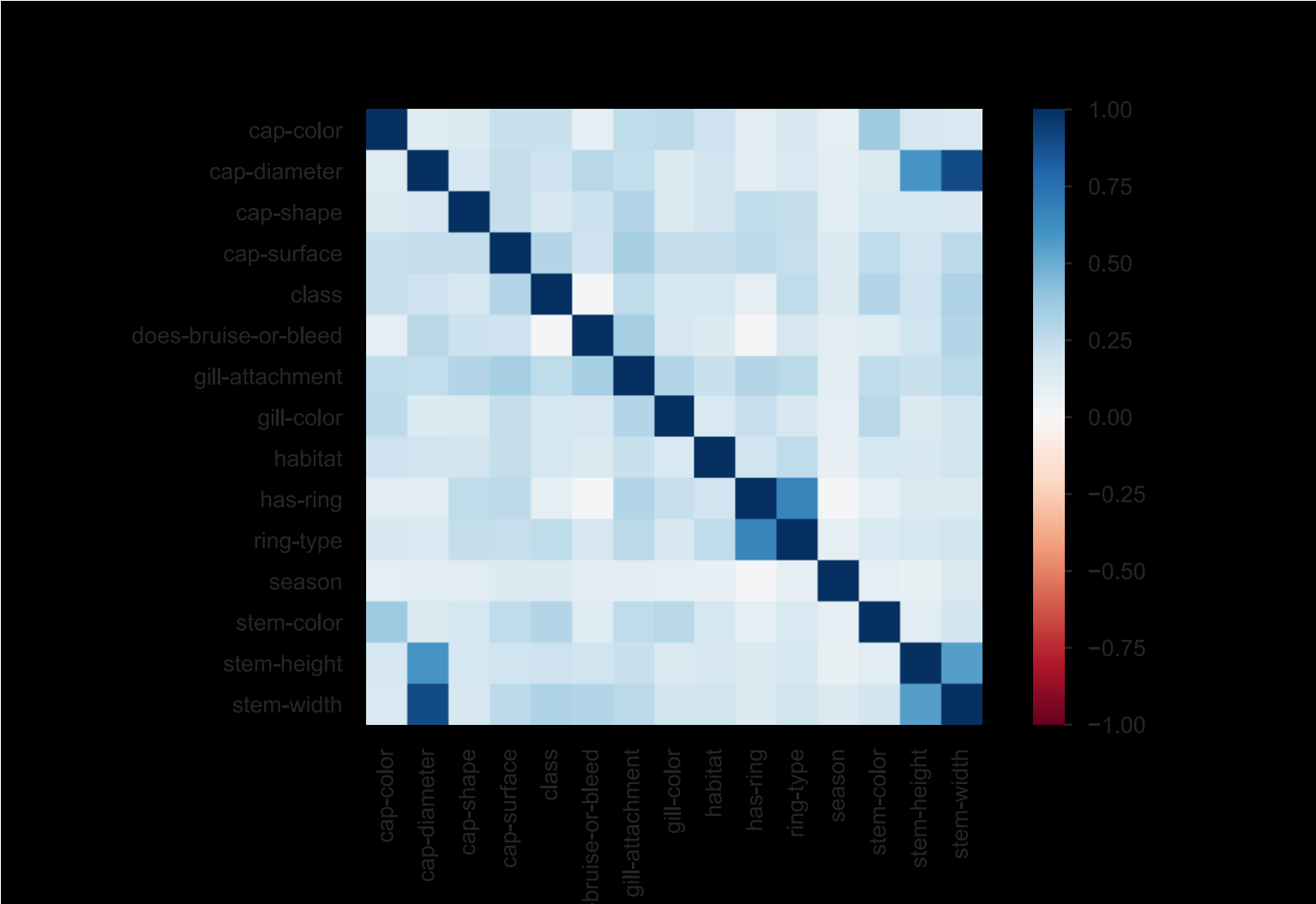


Interactions



Correlations

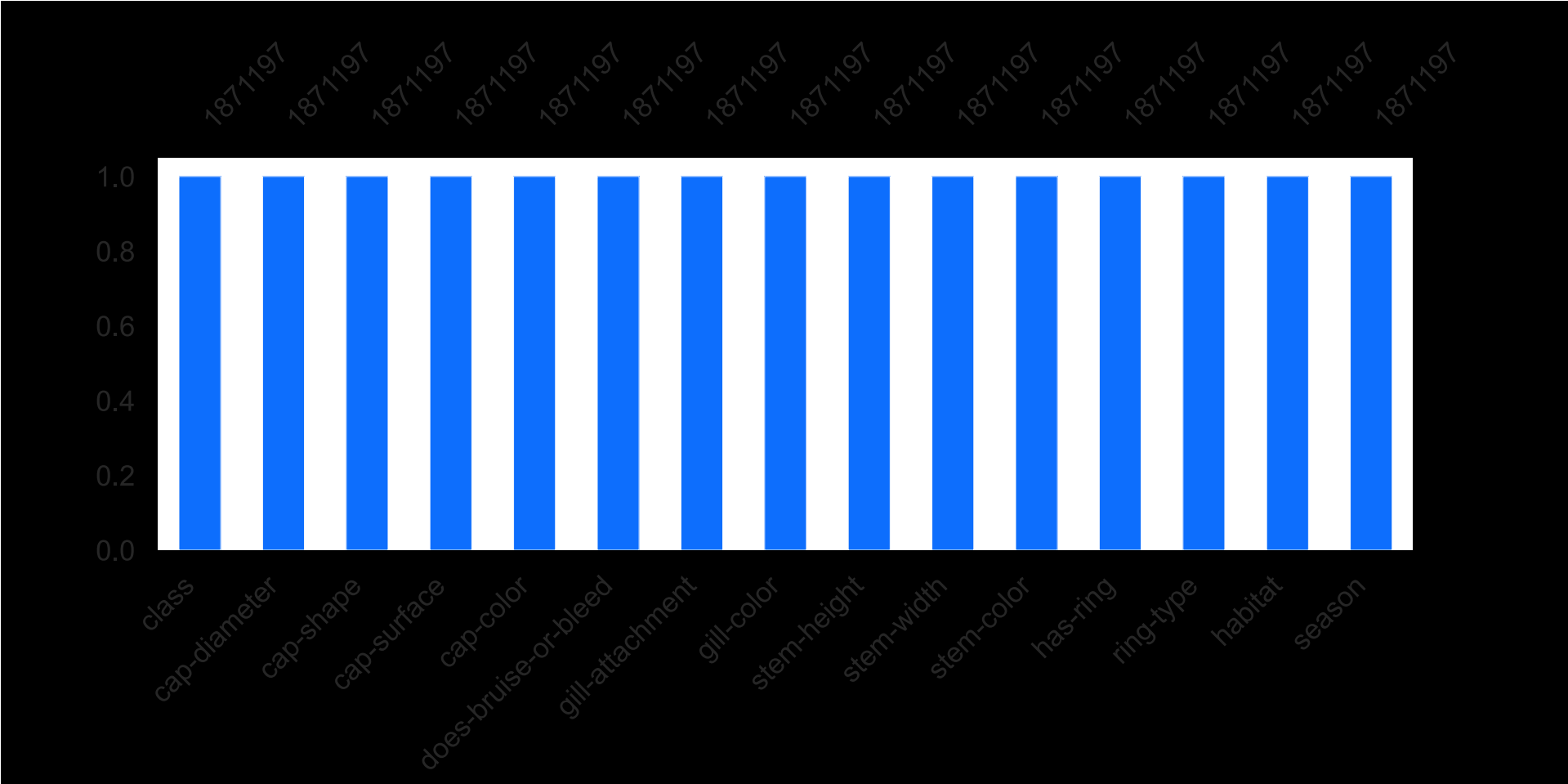




Missing values

Count

[Matrix](#)



A simple visualization of nullity by column.

Sample

First rows

Last rows

	class	cap-diameter	cap-shape	cap-surface	cap-color	does-bruise-or-bleed	gill-attachment	gill-color	stem-height	stem-width	stem-color	has-ring
0	e	8.80	f	s	Other	f	a	w	4.51	15.39	w	f
1	p	4.51	x	h	o	f	a	n	4.79	6.48	o	t
2	e	6.94	f	s	Other	f	x	w	6.85	9.93	n	f
3	e	3.88	f	y	g	f	s	g	4.16	6.53	w	f
4	e	5.85	x	Other	w	f	d	w	3.37	8.36	w	f
5	p	4.30	x	t	n	f	s	n	5.91	8.20	w	t

	class	cap-diameter	cap-shape	cap-surface	cap-color	does-bruise-or-bleed	gill-attachment	gill-color	stem-height	stem-width	stem-color	has-ring
7	p	4.55	x	e	e	f	a	y	8.31	9.77	y	t
8	p	7.36	f	h	e	f	x	w	5.77	17.13	w	f
9	e	6.45	x	t	n	f	a	w	7.13	12.77	e	f
10	e	7.49	f	t	n	f	s	n	9.14	13.00	o	f

Report generated by [YData](#).

