

# Overview

Brought to you by [YData](#)

Overview

Alerts 6

Reproduction

## Alerts

<a href="#">cap-diameter</a> is highly overall correlated with <a href="#">stem-height</a> and 1 other fields	High correlation
<a href="#">has-ring</a> is highly overall correlated with <a href="#">ring-type</a>	High correlation
<a href="#">ring-type</a> is highly overall correlated with <a href="#">has-ring</a>	High correlation
<a href="#">stem-height</a> is highly overall correlated with <a href="#">cap-diameter</a> and 1 other fields	High correlation
<a href="#">stem-width</a> is highly overall correlated with <a href="#">cap-diameter</a> and 1 other fields	High correlation
<a href="#">ring-type</a> is highly imbalanced (59.0%)	Imbalance

# Variables

Select Columns

[class](#)

Categorical

Distinct	2
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%

p

989989

e

828198

Memory size	15.6 MiB
-------------	----------

More details

- Overview
- Categories
- Words
- Characters

Common Values

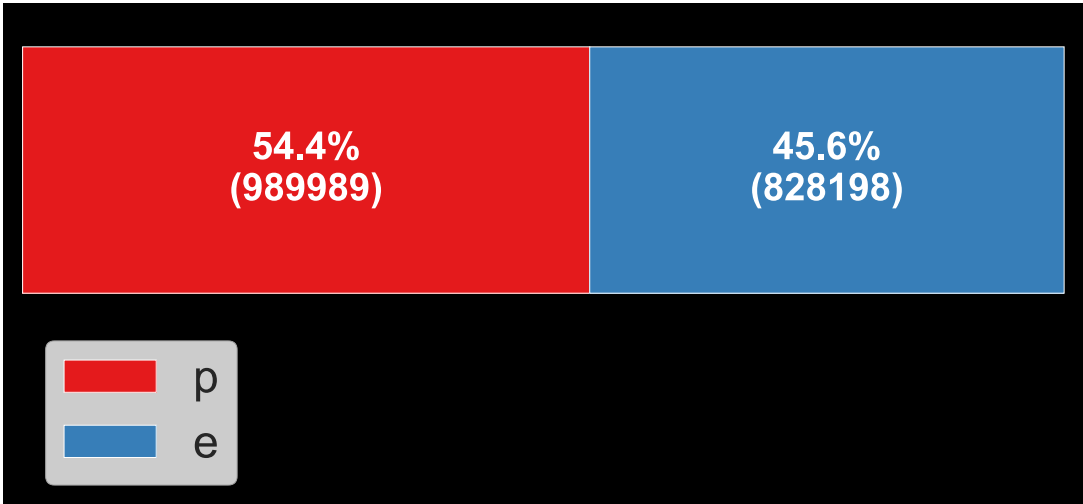
Value	Count	Frequency (%)
p	989989	54.4%
e	828198	45.6%

Length



Histogram of lengths of the category

Common Values (Plot)

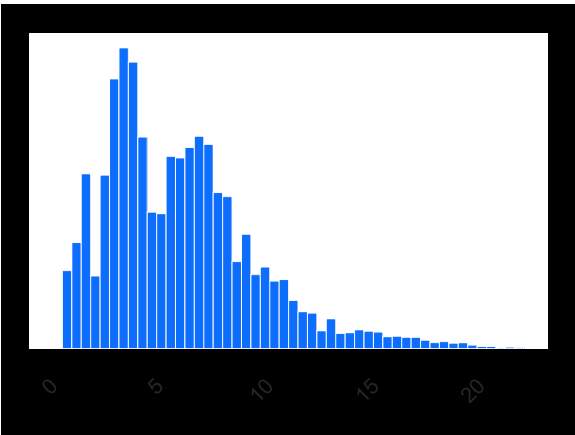


cap-diameter

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	6.1450704	Memory size	27.7 MiB



More details

Statistics

Histogram

Common values

Extreme values

Quantile statistics

Minimum	0.1
5-th percentile	1.46
Q1	3.44

Descriptive statistics

Standard deviation	3.5953521
Coefficient of variation (CV)	0.58507907
Kurtosis	1.4000363

median	5.71	Mean	6.1450704
Q3	8.03	Median Absolute Deviation (MAD)	2.29
95-th percentile	13.05	Skewness	1.0618077
Maximum	22.38	Sum	11172887
Range	22.28	Variance	12.926557
Interquartile range (IQR)	4.59	Monotonicity	Not monotonic

cap-shape

Categorical

Distinct	6
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	15.6 MiB

x	949087
f	411120
s	227478
b	127391
c	55767

More details

Overview

Categories

Words

Characters

Common Values

Value	Count	Frequency (%)
x	949087	52.2%
f	411120	22.6%
s	227478	12.5%
b	127391	7.0%

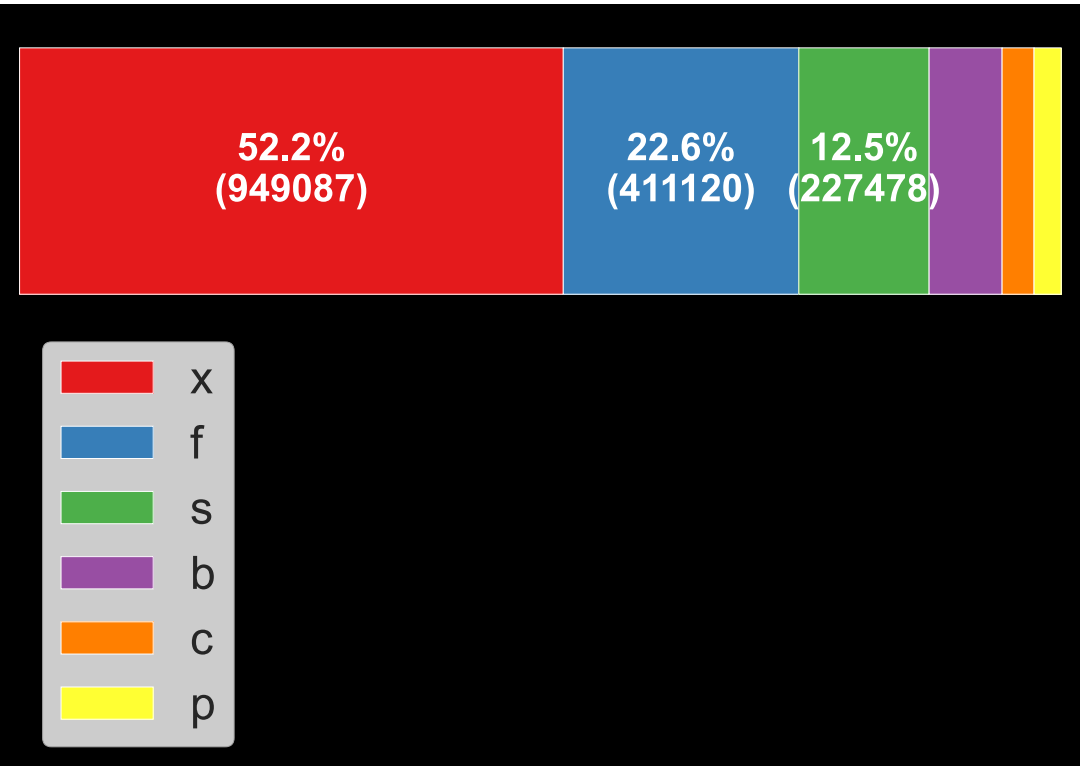
Value	Count	Frequency (%)
c	55767	3.1%
p	47344	2.6%

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	15.6 MiB

t	403314
s	317405
h	208798
g	193750
d	193381
Other values ...	501539

More details

Overview

Categories

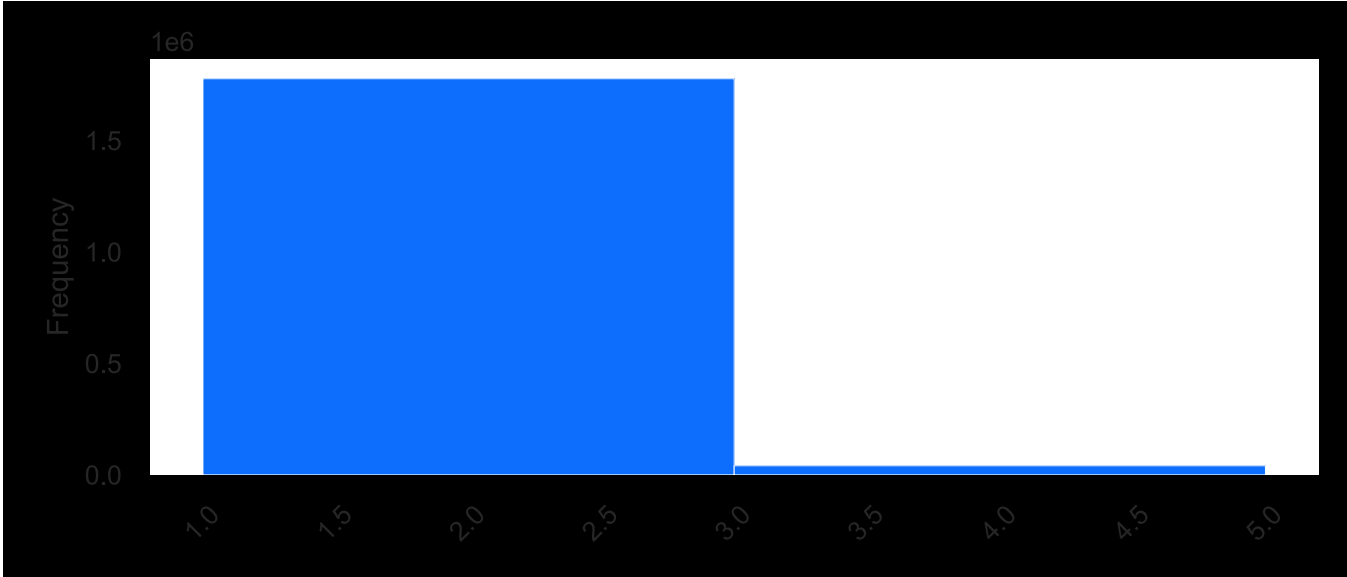
Words

Characters

Common Values

Value	Count	Frequency (%)
t	403314	22.2%
s	317405	17.5%
h	208798	11.5%
g	193750	10.7%
d	193381	10.6%
y	155516	8.6%
e	87269	4.8%
k	84501	4.6%
w	70730	3.9%
i	62626	3.4%

Length



Histogram of lengths of the category

cap-color

Categorical

Distinct	8
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	15.6 MiB

n	834231
w	224959
Other	181806
y	166820
e	126695
Other values ...	283676

More details

Overview

Categories

Words

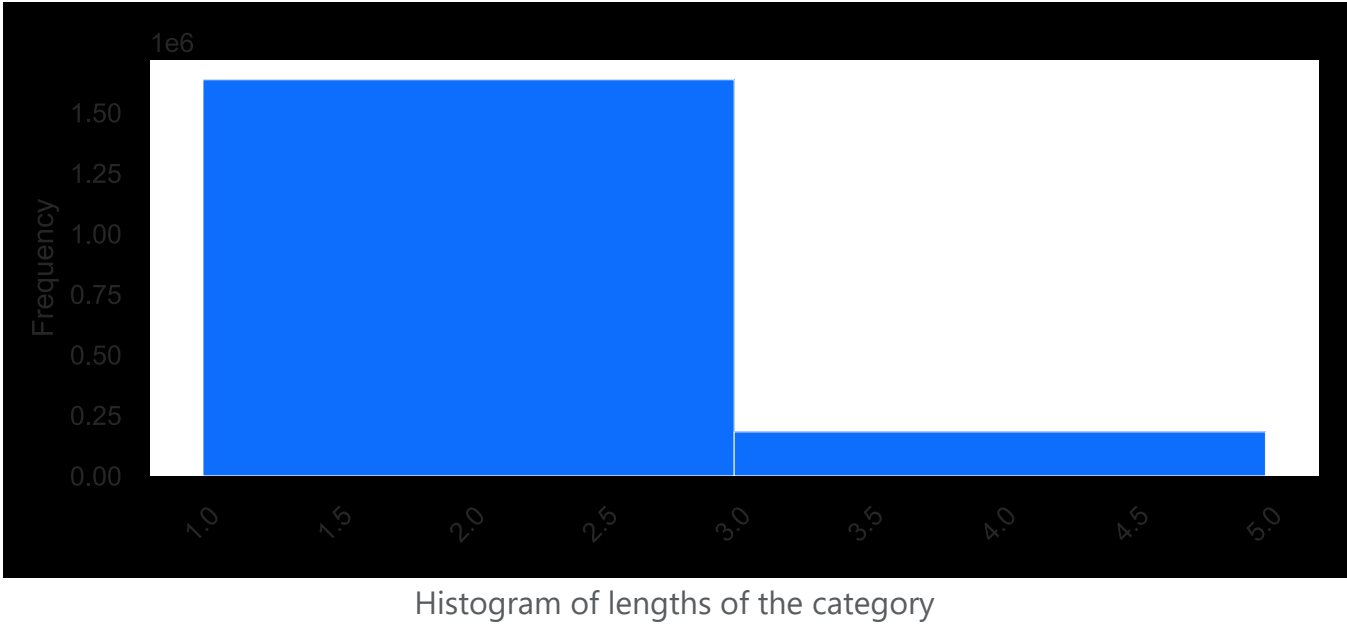
Characters

Common Values

Value	Count	Frequency (%)
n	834231	45.9%
w	224959	12.4%
Other	181806	10.0%

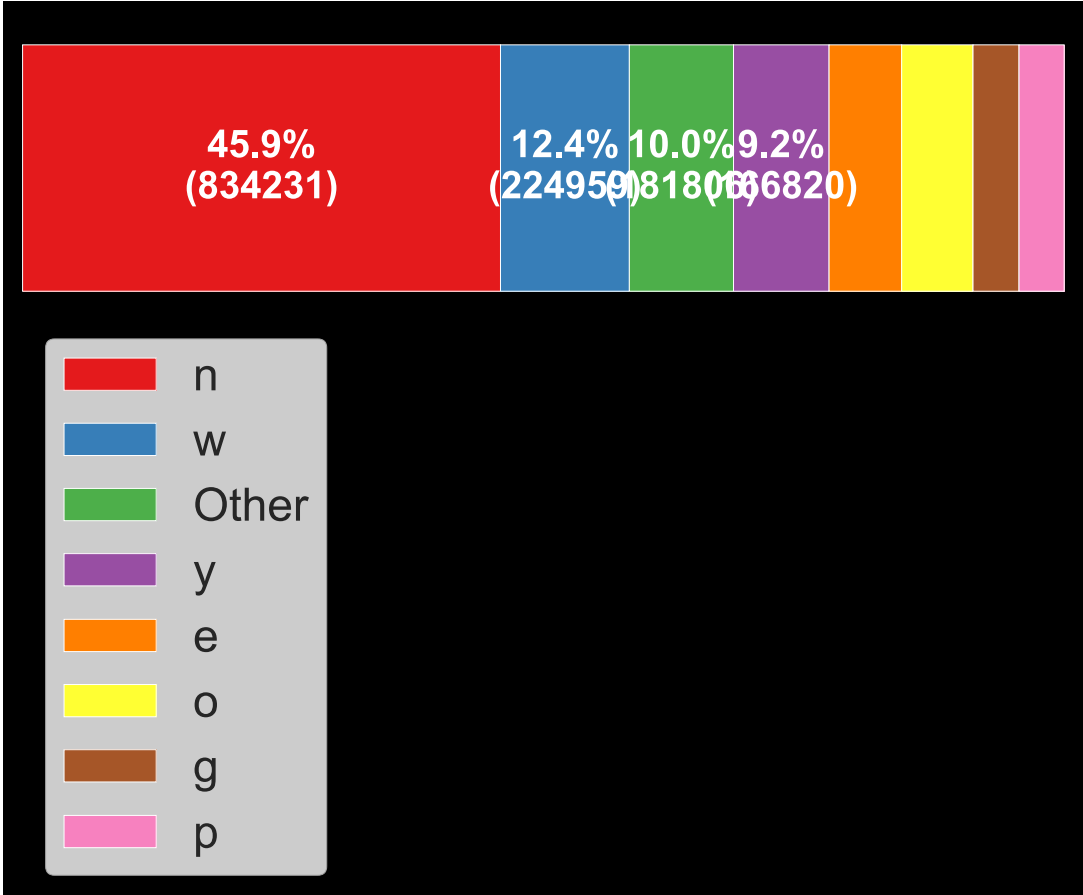
Value	Count	Frequency (%)
y	166820	9.2%
e	126695	7.0%
o	124576	6.9%
g	80297	4.4%
p	78803	4.3%

Length



Common Values (Plot)





does-bruise-or-bleed

Categorical

Distinct	2
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	15.6 MiB

f 1418777  
t 399410

More details

OverviewCategoriesWordsCharacters

Common Values

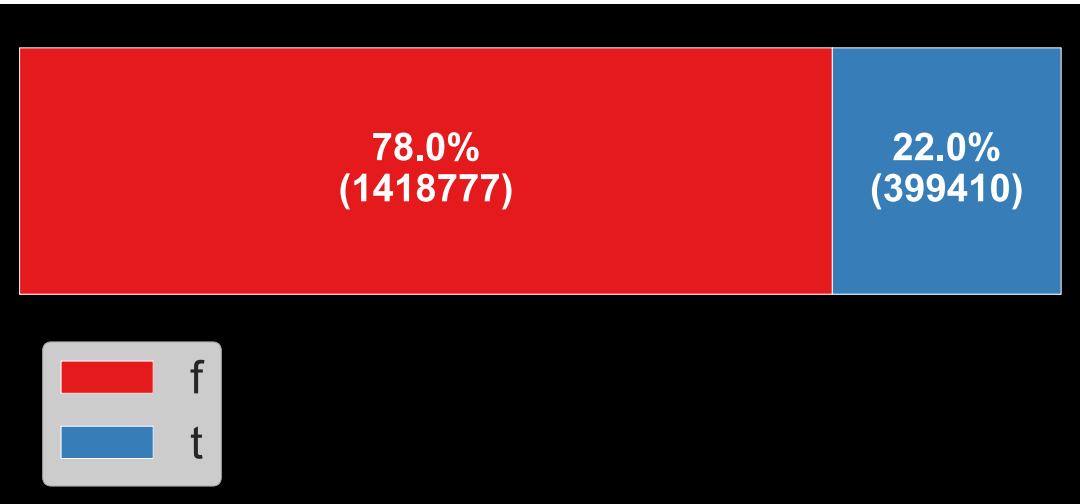
Value	Count	Frequency (%)
f	1418777	78.0%
t	399410	22.0%

Length



Histogram of lengths of the category

Common Values (Plot)



**gill-attachment**

Categorical

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

a

510772

Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	15.6 MiB

d	417115
x	254121
s	241547
p	203079

More details

Overview

Categories

Words

Characters

Common Values

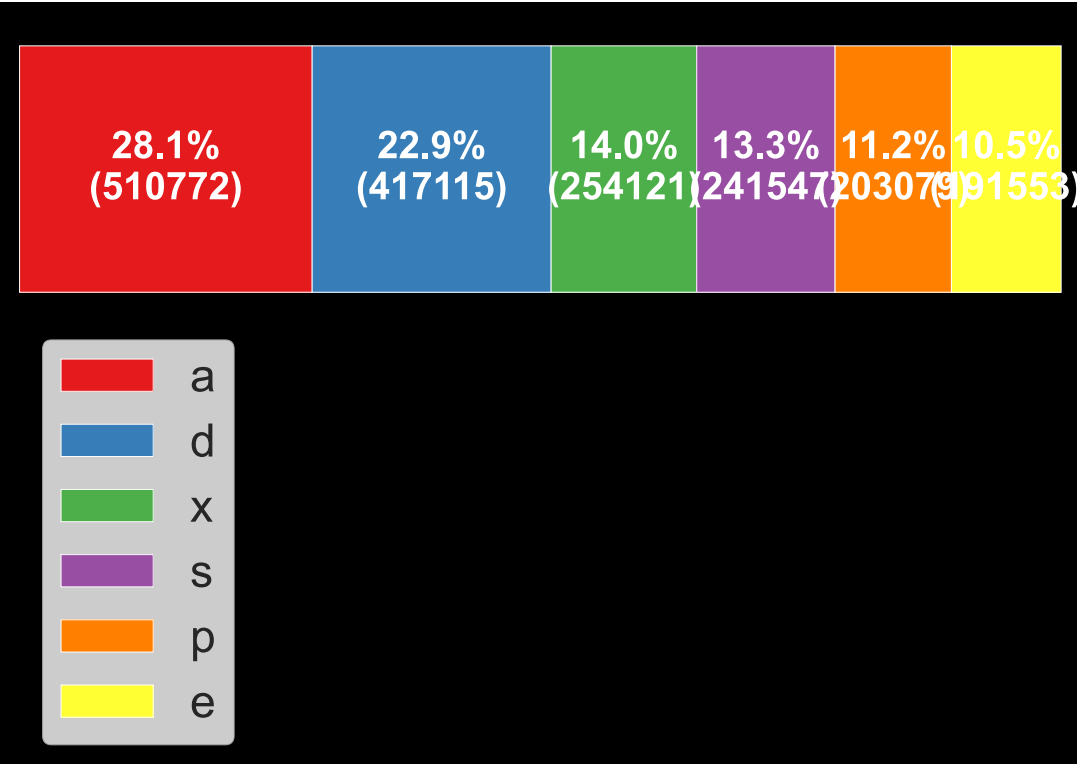
Value	Count	Frequency (%)
a	510772	28.1%
d	417115	22.9%
x	254121	14.0%
s	241547	13.3%
p	203079	11.2%
e	191553	10.5%

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	15.6 MiB

w	561649
n	354308
y	313718
p	266476
Other	116311
Other values ...	205725

More details

Overview

Categories

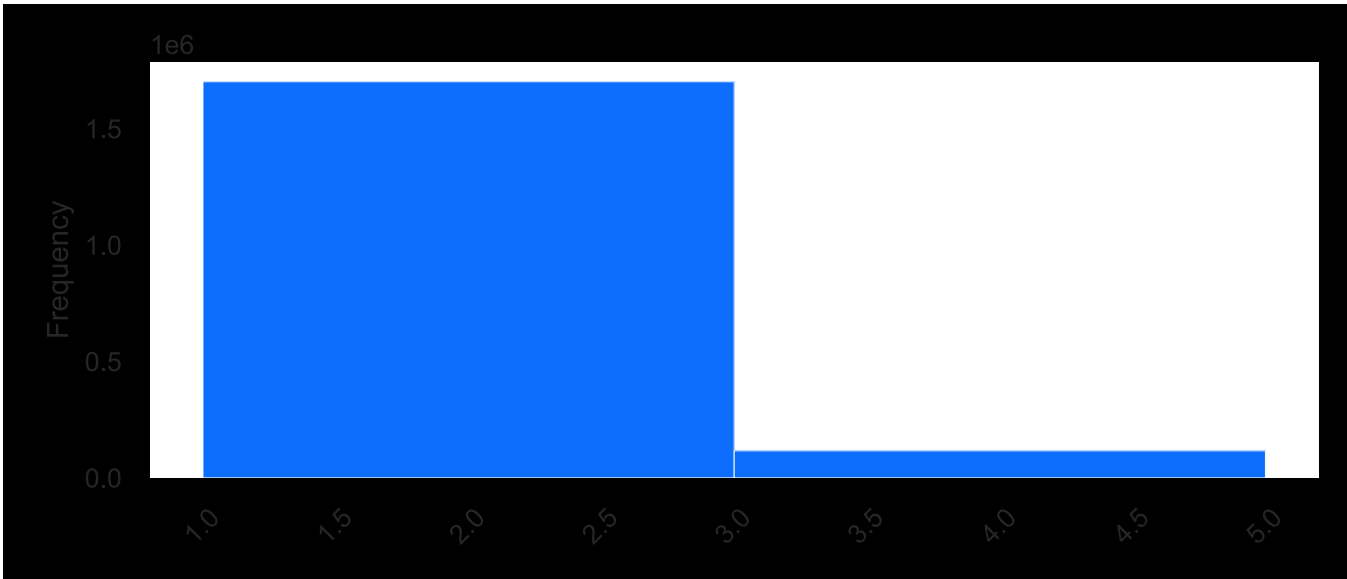
Words

Characters

Common Values

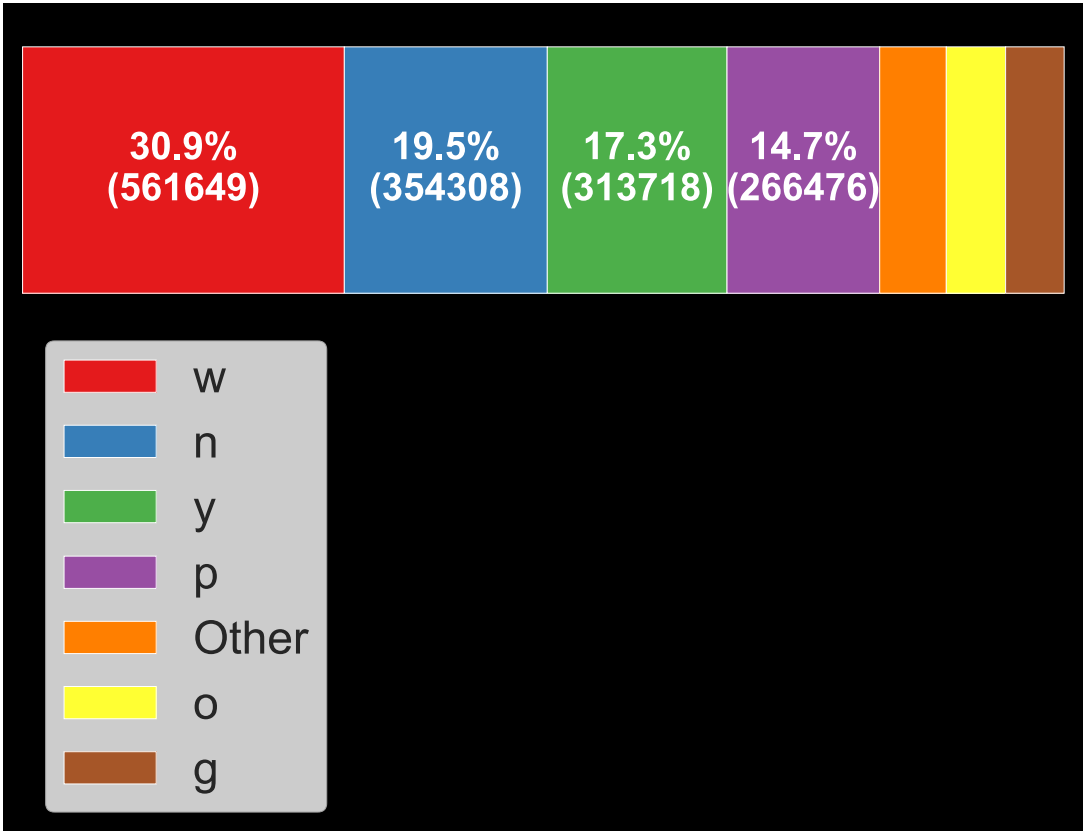
Value	Count	Frequency (%)
w	561649	30.9%
n	354308	19.5%
y	313718	17.3%
p	266476	14.7%
Other	116311	6.4%
o	103490	5.7%
g	102235	5.6%

Length



Histogram of lengths of the category

Common Values (Plot)

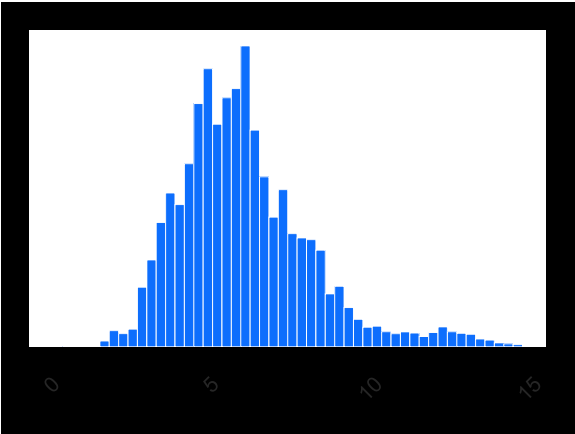


stem-height

Real number ( $\mathbb{R}$ )

High correlation

Distinct	1351	Minimum	0
Distinct (%)	0.1%	Maximum	14.75
Missing	0	Zeros	1
Missing (%)	0.0%	Zeros (%)	< 0.1%
Infinite	0	Negative	0
Infinite (%)	0.0%	Negative (%)	0.0%
Mean	6.0215922	Memory size	27.7 MiB



More details

Statistics

Histogram

Common values

Extreme values

Quantile statistics

Minimum	0
5-th percentile	3.22
Q1	4.61
median	5.75
Q3	7.09
95-th percentile	9.96
Maximum	14.75
Range	14.75
Interquartile range (IQR)	2.48

Descriptive statistics

Standard deviation	2.0948512
Coefficient of variation (CV)	0.34788992
Kurtosis	1.6636055
Mean	6.0215922
Median Absolute Deviation (MAD)	1.21
Skewness	1.05431
Sum	10948381
Variance	4.3884016
Monotonicity	Not monotonic

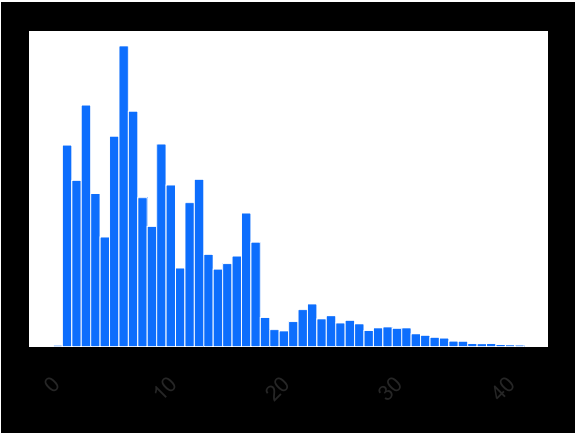
stem-width

Real number ( $\mathbb{R}$ )

High correlation

Distinct	4018
Distinct (%)	0.2%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Mean	10.61804

Minimum	0
Maximum	41.8
Zeros	1
Zeros (%)	< 0.1%
Negative	0
Negative (%)	0.0%
Memory size	27.7 MiB



More details

Statistics

Histogram

Common values

Extreme values

Quantile statistics

Minimum	0
5-th percentile	1.61
Q1	5.28
median	8.85
Q3	14.45
95-th percentile	26.73
Maximum	41.8
Range	41.8
Interquartile range (IQR)	9.17

Descriptive statistics

Standard deviation	7.6162992
Coefficient of variation (CV)	0.71729804
Kurtosis	1.1842502
Mean	10.61804
Median Absolute Deviation (MAD)	4.41
Skewness	1.17239
Sum	19305582
Variance	58.008014
Monotonicity	Not monotonic

stem-color

Categorical

Distinct	7
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	15.6 MiB

n	659046
w	590588
y	184812
Other	122768
o	96065
Other values ...	164908

More details

Overview

Categories

Words

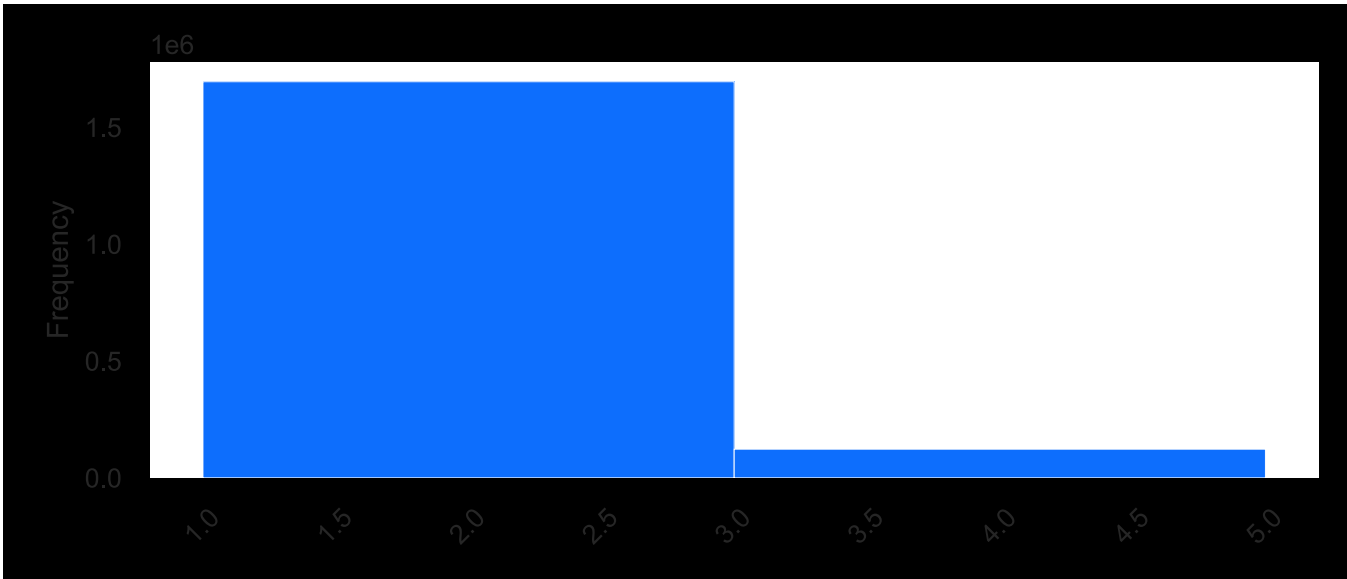
Characters

Common Values



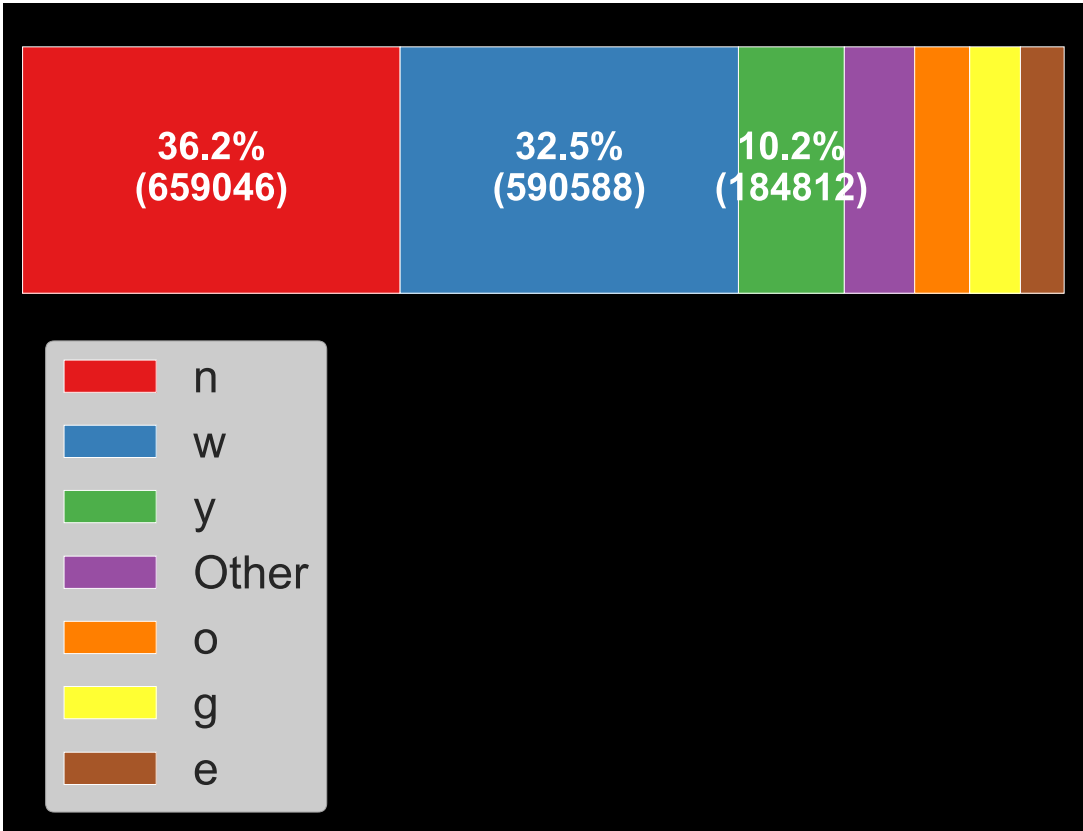
Value	Count	Frequency (%)
n	659046	36.2%
w	590588	32.5%
y	184812	10.2%
Other	122768	6.8%
o	96065	5.3%
g	88576	4.9%
e	76332	4.2%

Length



Histogram of lengths of the category

Common Values (Plot)



has-ring

Categorical

High correlation

Distinct	2
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	15.6 MiB

f 1465417  
t 352770

More details

OverviewCategoriesWordsCharacters

Common Values

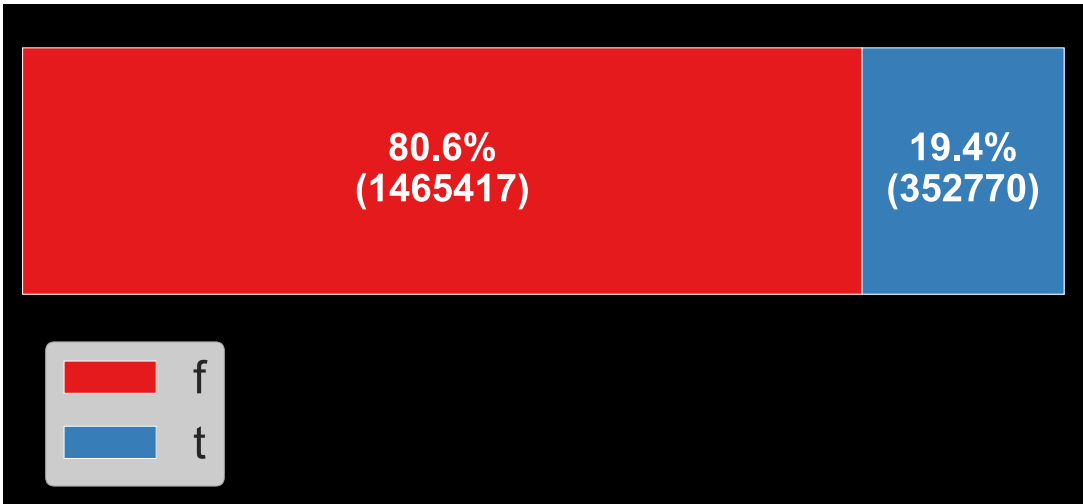
Value	Count	Frequency (%)
f	1465417	80.6%
t	352770	19.4%

Length



Histogram of lengths of the category

Common Values (Plot)



ring-type

Categorical

High correlation Imbalance

Distinct	6
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	15.6 MiB

f	1503127
z	80788
e	78942
r	59037
p	50633

More details

Overview

Categories

Words

Characters

Common Values

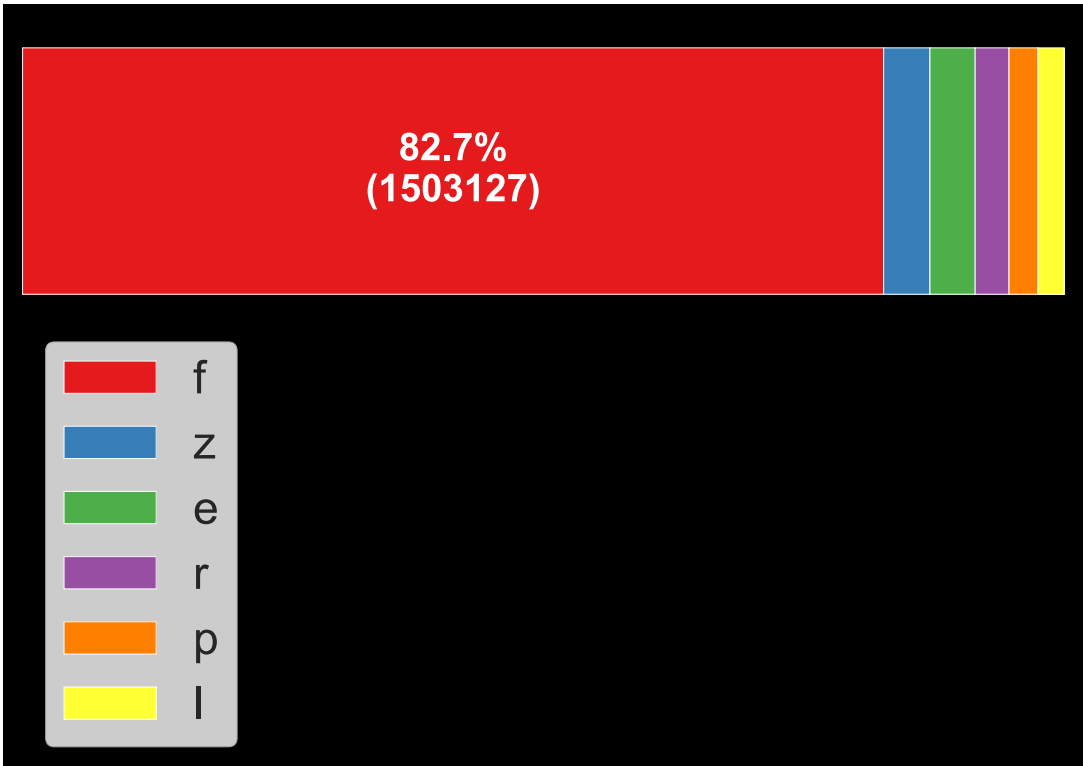
Value	Count	Frequency (%)
f	1503127	82.7%
z	80788	4.4%
e	78942	4.3%
r	59037	3.2%
p	50633	2.8%
l	45660	2.5%

Length



Histogram of lengths of the category

Common Values (Plot)



habitat

Categorical

Distinct	5
Distinct (%)	< 0.1%

d	1218018
g	260747
m	127205

Missing	0
Missing (%)	0.0%
Memory size	15.6 MiB

l 110484  
h 101733

More details

Overview

Categories

Words

Characters

Length

Max length	1
Median length	1
Mean length	1
Min length	1

Characters and Unicode

Total characters	1818187	
Distinct characters	5	
Distinct categories	1	?
Distinct scripts	1	?
Distinct blocks	1	?

The Unicode Standard assigns character properties to each code point, which can be used to analyse textual variables.

Unique

Unique	0	?
Unique (%)	0.0%	

Sample

1st row	d
2nd row	d
3rd row	l
4th row	d
5th row	g

season

Categorical

Distinct	4
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	15.6 MiB

a 914618  
u 693002  
w 159200  
s 51367

More details

Overview

Categories

Words

Characters

Common Values

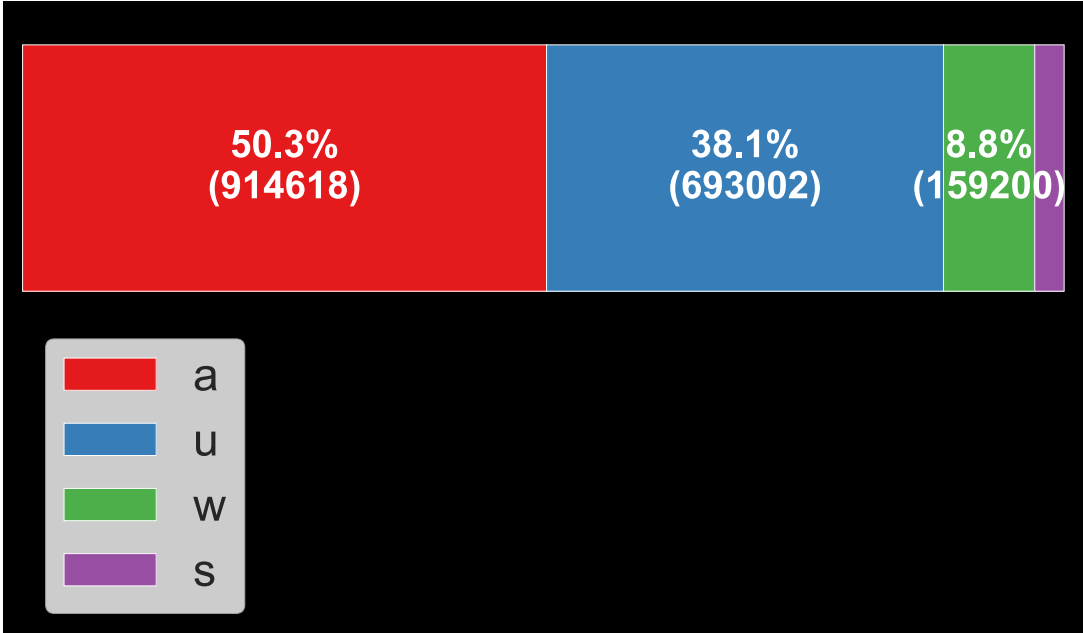
Value	Count	Frequency (%)
a	914618	50.3%
u	693002	38.1%
w	159200	8.8%
s	51367	2.8%

Length



Histogram of lengths of the category

Common Values (Plot)



# Interactions

cap-diameter

stem-height

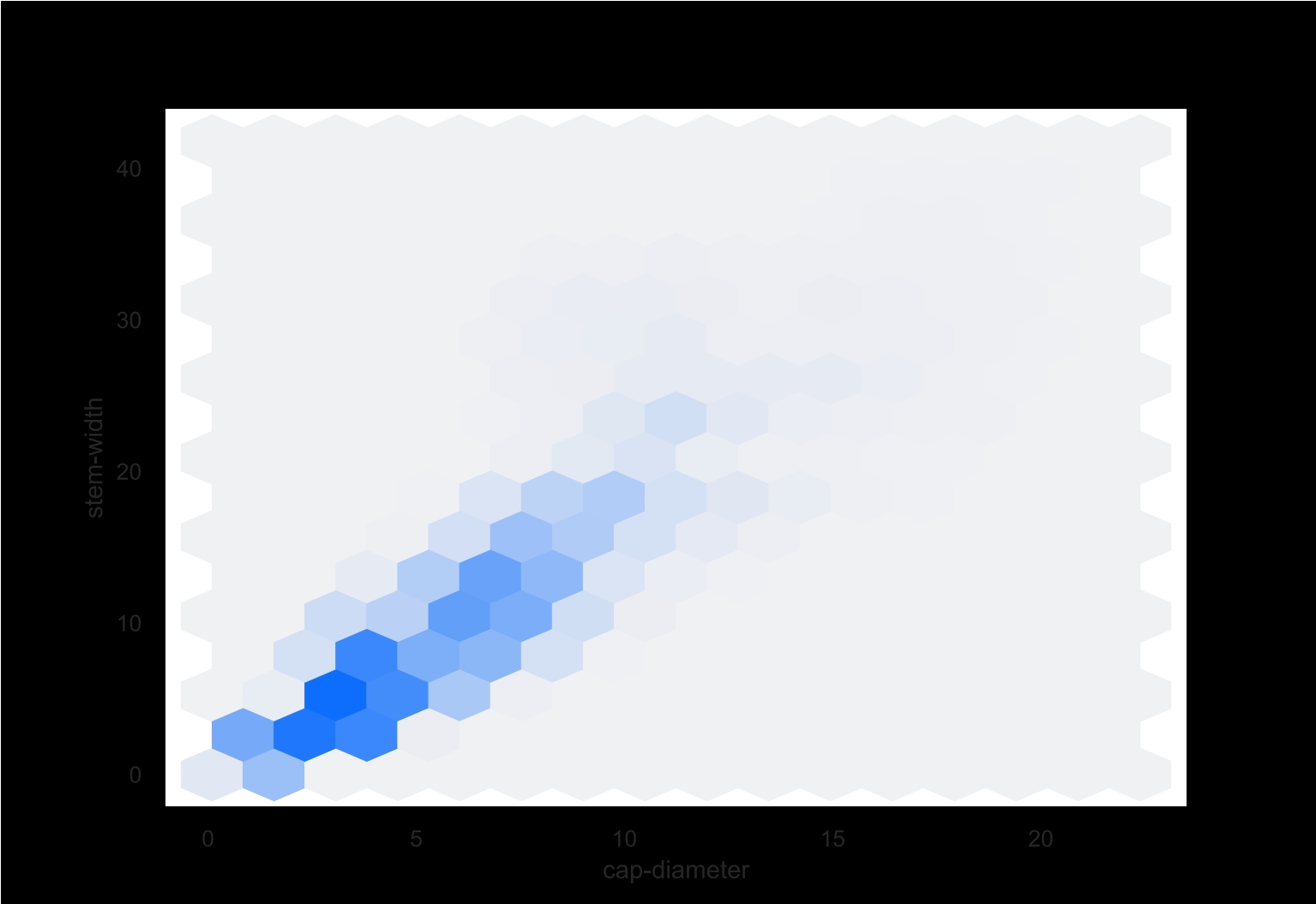
stem-width

stem-width

cap-diameter

stem-height



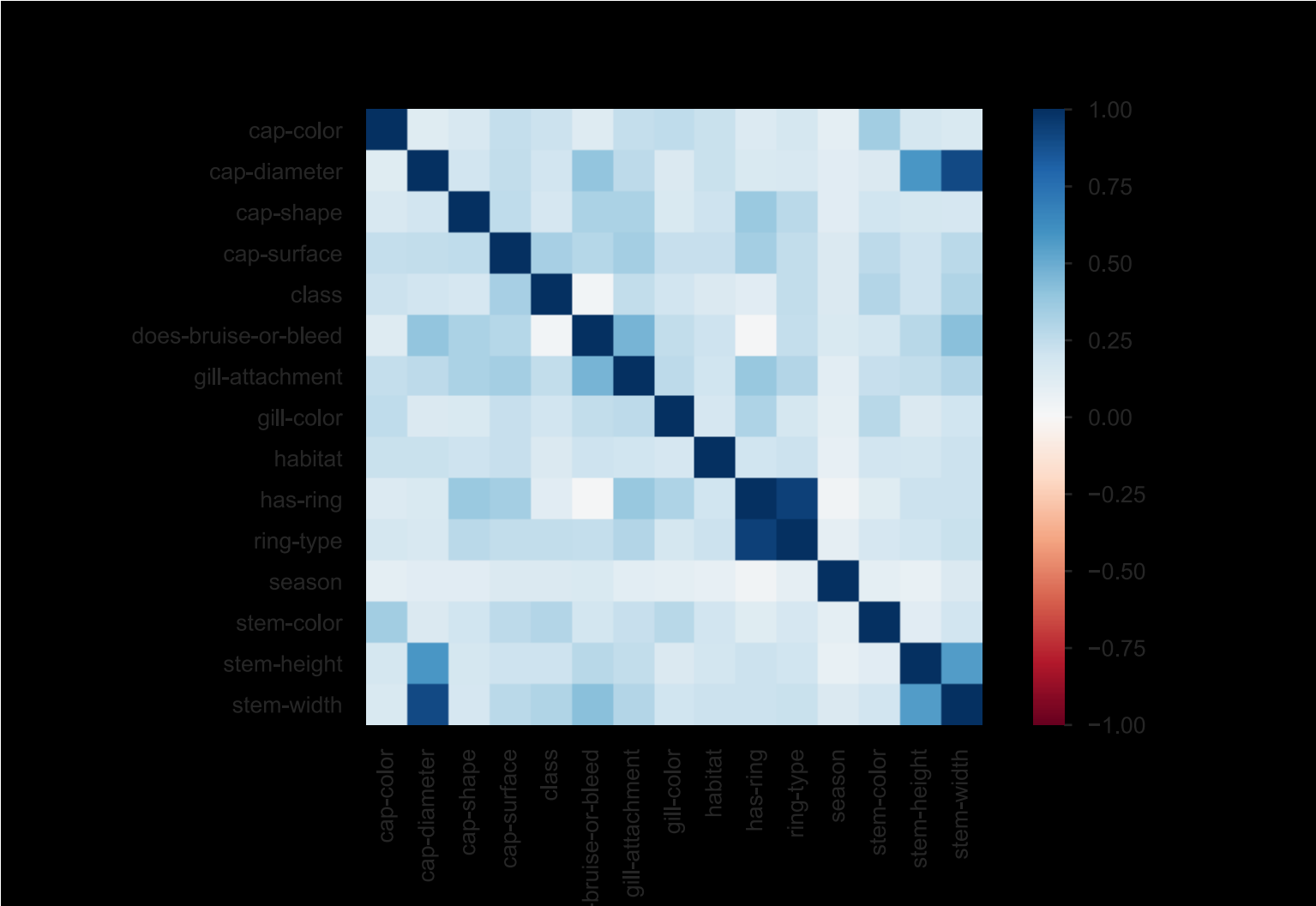


# Correlations

Auto

Heatmap

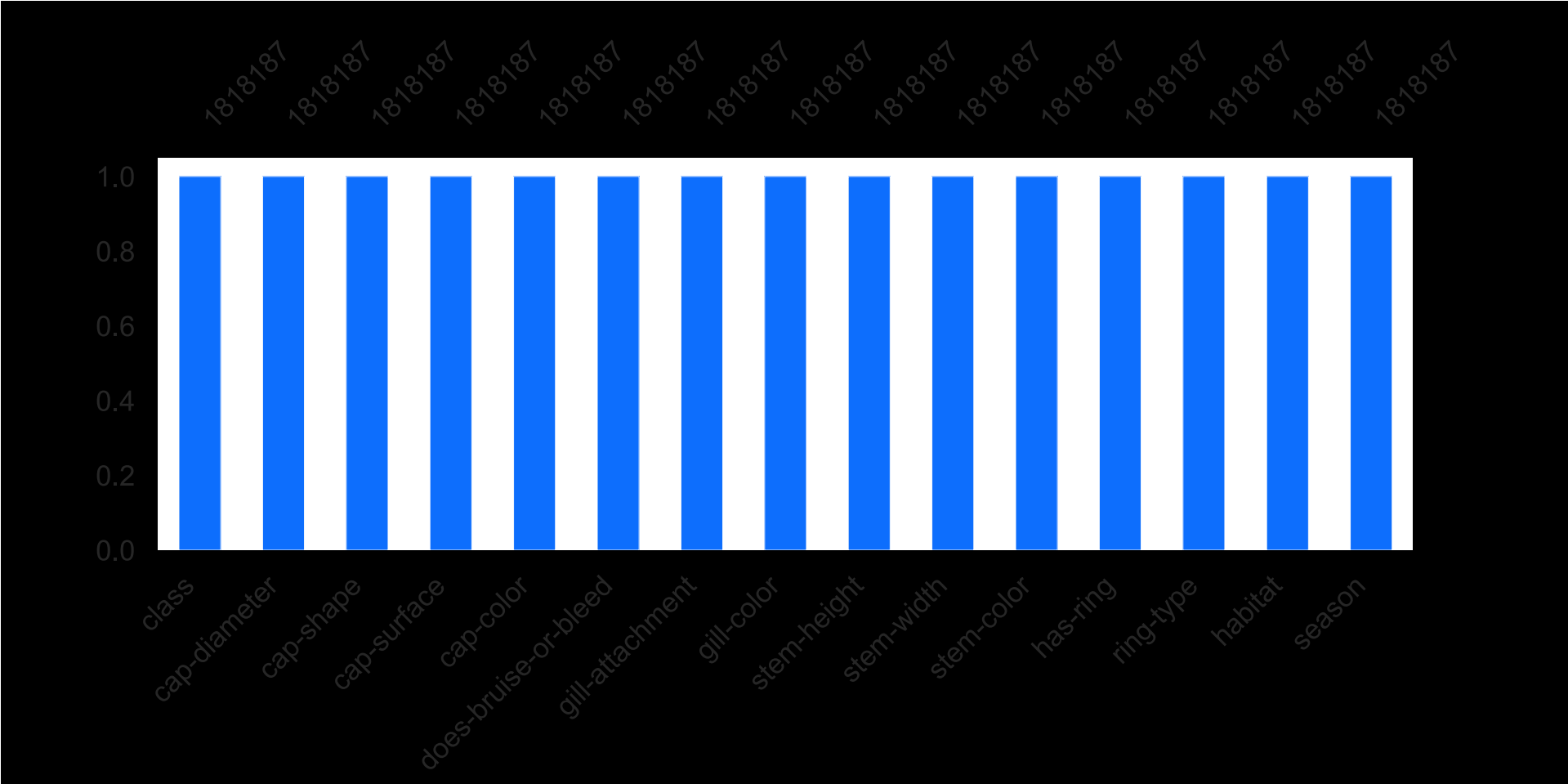
Table



# 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](#).

