

Overview

Dataset statistics	
Number of variables	11
Number of observations	193573
Missing cells	0
Missing cells (%)	0.0%
Duplicate rows	0
Duplicate rows (%)	0.0%
Total size in memory	45.3 MiB
Average record size in memory	245.3 B

Variable types	
Numeric	8
Categorical	3


Alerts

carat is highly overall correlated with x and 3 other fields (x, y, z, price)	High correlation
x is highly overall correlated with carat and 3 other fields (carat, y, z, price)	High correlation
y is highly overall correlated with carat and 3 other fields (carat, x, z, price)	High correlation
z is highly overall correlated with carat and 3 other fields (carat, x, y, price)	High correlation
price is highly overall correlated with carat and 3 other fields (carat, x, y, z)	High correlation
id is uniformly distributed	Uniform
id has unique values	Unique

Reproduction

Analysis started	2023-07-28 05:21:05.169256
Analysis finished	2023-07-28 05:21:33.668925
Duration	28.5 seconds
Software version	pandas-profiling v3.6.6 (https://github.com/pandas-profiling/pandas-profiling)
Download configuration	config.json (data:text/plain;charset=utf-8,%7B%22title%22%3A%20%22Profiling%20Report%22%2C%20%22dataset%22%3A%20%7B%22description%22%3A%20%22%22%2C%20%22create

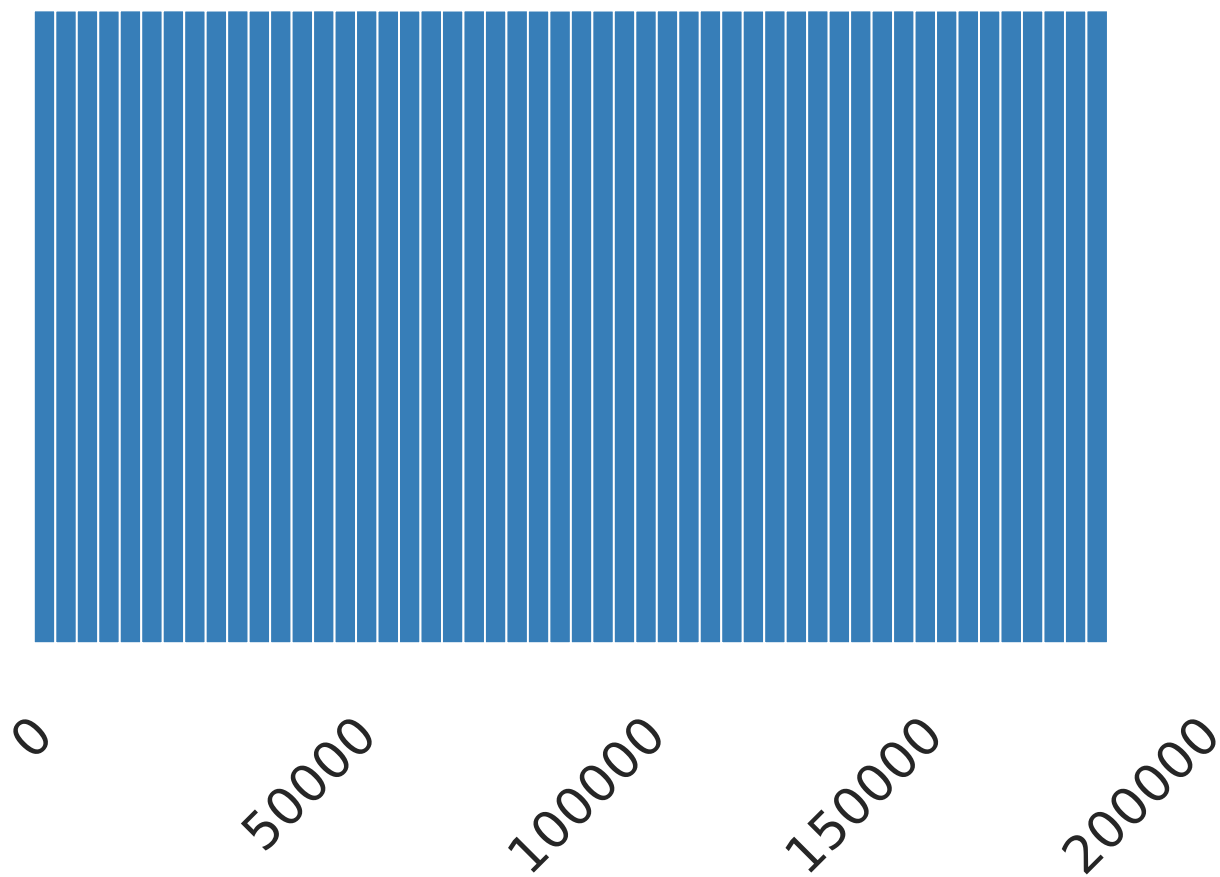
Variables

Select Columns 

id
Real number (ℝ)

UNIFORM UNIQUE

Distinct	193573
Distinct (%)	100.0%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Mean	96786
Minimum	0
Maximum	193572
Zeros	1
Zeros (%)	< 0.1%
Negative	0
Negative (%)	0.0%
Memory size	1.5 MiB



Quantile statistics

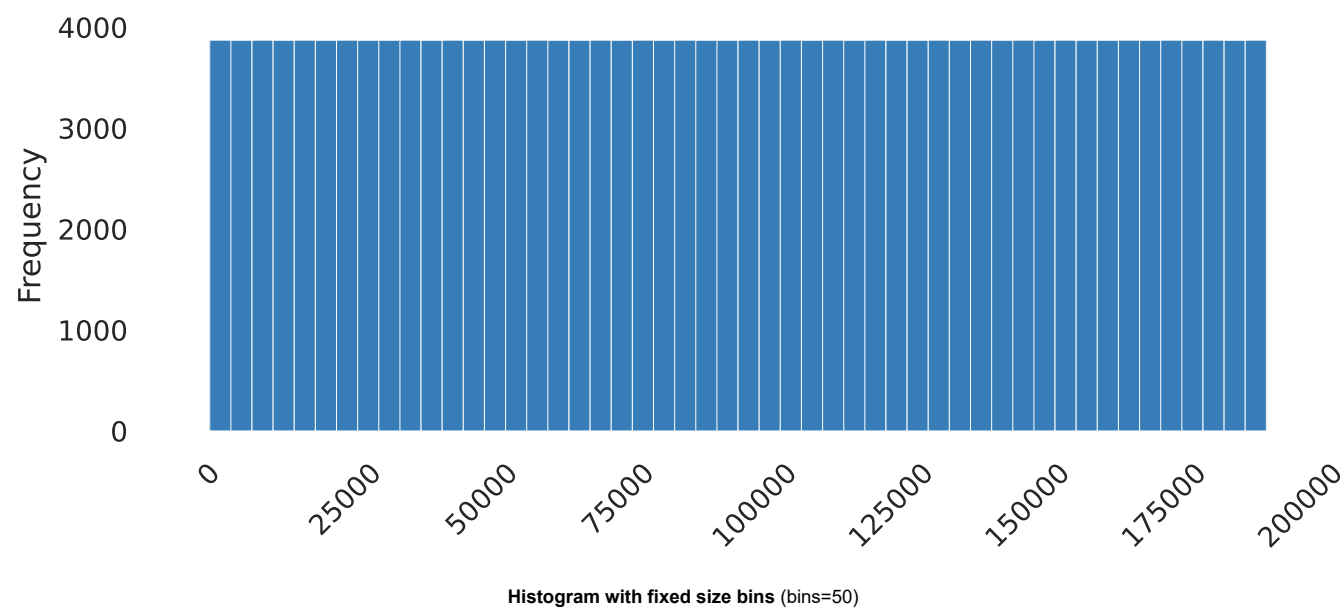
Minimum	0
5-th percentile	9678.6
Q1	48393
median	96786
Q3	145179
95-th percentile	183893.4
Maximum	193572
Range	193572
Interquartile range (IQR)	96786

Descriptive statistics

Standard deviation	55879.856
Coefficient of variation (CV)	0.57735474
Kurtosis	-1.2
Mean	96786
Median Absolute Deviation (MAD)	48393
Skewness	0
Sum	1.8735156×10^{10}
Variance	3.1225583×10^9

Monotonicity

Strictly increasing



Value	Count	Frequency (%)
0	1	< 0.1%
129053	1	< 0.1%
129044	1	< 0.1%
129045	1	< 0.1%
129046	1	< 0.1%
129047	1	< 0.1%
129048	1	< 0.1%
129049	1	< 0.1%
129050	1	< 0.1%
129051	1	< 0.1%
Other values (193563)	193563	> 99.9%

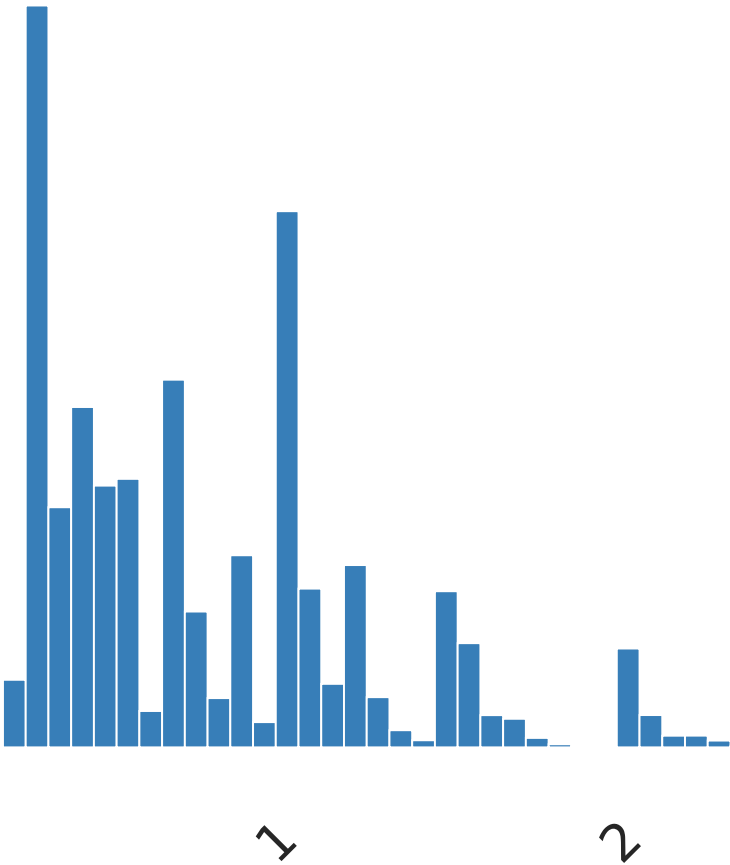
Value	Count	Frequency (%)
0	1	< 0.1%
1	1	< 0.1%
2	1	< 0.1%
3	1	< 0.1%
4	1	< 0.1%
5	1	< 0.1%
6	1	< 0.1%
7	1	< 0.1%
8	1	< 0.1%
9	1	< 0.1%

Value	Count	Frequency (%)
193572	1	< 0.1%
193571	1	< 0.1%
193570	1	< 0.1%
193569	1	< 0.1%
193568	1	< 0.1%
193567	1	< 0.1%
193566	1	< 0.1%
193565	1	< 0.1%
193564	1	< 0.1%
193563	1	< 0.1%

carat

Real number (R)

Distinct	248
Distinct (%)	0.1%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Mean	0.79068785
Minimum	0.2
Maximum	3.5
Zeros	0
Zeros (%)	0.0%
Negative	0
Negative (%)	0.0%
Memory size	1.5 MiB

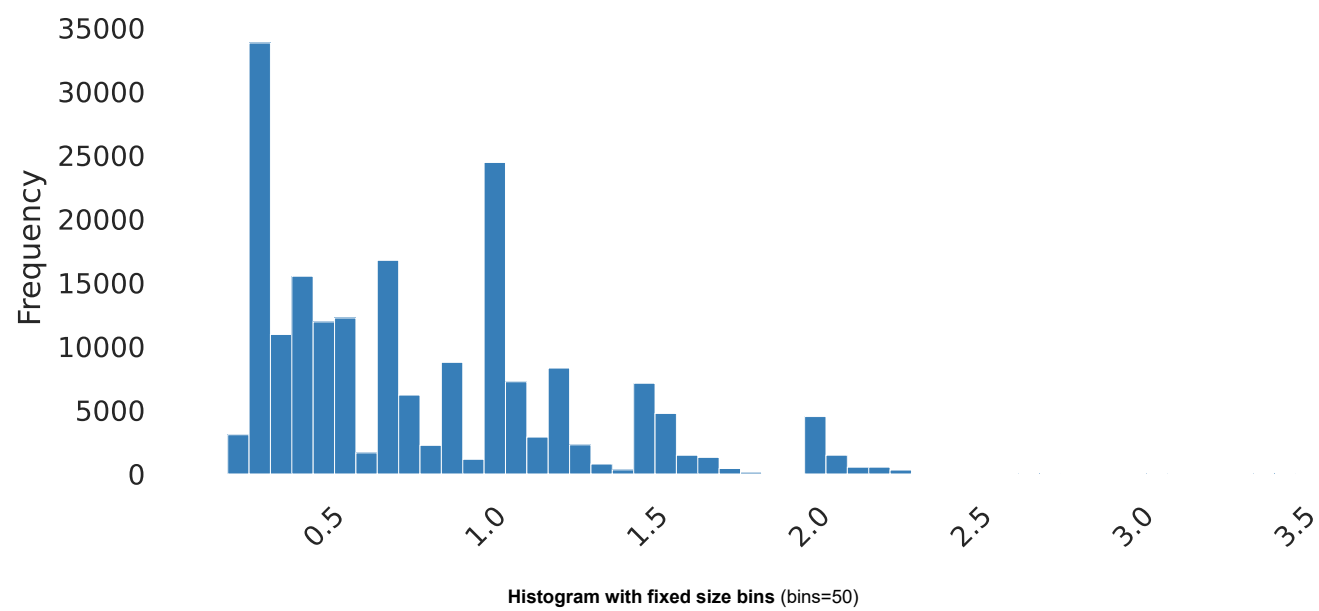


Quantile statistics

Minimum	0.2
5-th percentile	0.3
Q1	0.4
median	0.7
Q3	1.03
95-th percentile	1.65
Maximum	3.5
Range	3.3
Interquartile range (IQR)	0.63

Descriptive statistics

Standard deviation	0.46268774
Coefficient of variation (CV)	0.58517117
Kurtosis	0.53739769
Mean	0.79068785
Median Absolute Deviation (MAD)	0.32
Skewness	0.99513461
Sum	153055.82
Variance	0.21407994
Monotonicity	Not monotonic



Value	Count	Frequency (%)
0.3	10758	5.6%
1.01	10103	5.2%
0.31	9538	4.9%
0.7	7958	4.1%
0.32	7548	3.9%
0.9	6253	3.2%
0.41	5852	3.0%
0.71	5367	2.8%
1	5328	2.8%
0.4	4802	2.5%
Other values (238)	120066	62.0%

Value	Count	Frequency (%)
0.2	30	< 0.1%
0.21	17	< 0.1%
0.22	3	< 0.1%
0.23	889	0.5%
0.24	809	0.4%
0.25	611	0.3%
0.26	722	0.4%
0.27	650	0.3%
0.28	466	0.2%
0.29	354	0.2%

Value	Count	Frequency (%)
3.5	1	< 0.1%
3.4	1	< 0.1%
3.04	3	< 0.1%
3.01	7	< 0.1%
3	5	< 0.1%
2.74	3	< 0.1%
2.72	1	< 0.1%
2.71	1	< 0.1%
2.7	2	< 0.1%
2.66	3	< 0.1%

cut
Categorical

Distinct	5
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	11.7 MiB

Length

Max length	9
Median length	7
Mean length	6.2214565
Min length	4

Characters and Unicode

Total characters	1204306	
Distinct characters	16	
Distinct categories	3 (https://en.wikipedia.org/wiki/Unicode_character_property#General_Category)	?
Distinct scripts	2 (https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode)	?
Distinct blocks	1 (https://en.wikipedia.org/wiki/Unicode_block)	?

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

Unique

Unique	0	?
Unique (%)	0.0%	

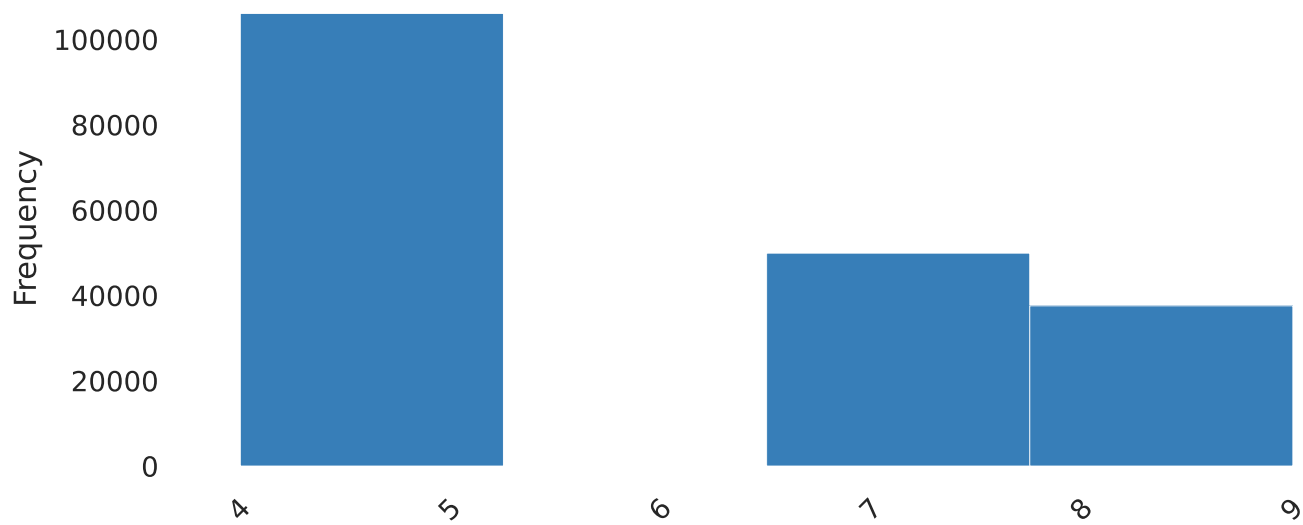
Sample

1st row	Premium
2nd row	Very Good
3rd row	Ideal
4th row	Ideal
5th row	Premium

Common Values

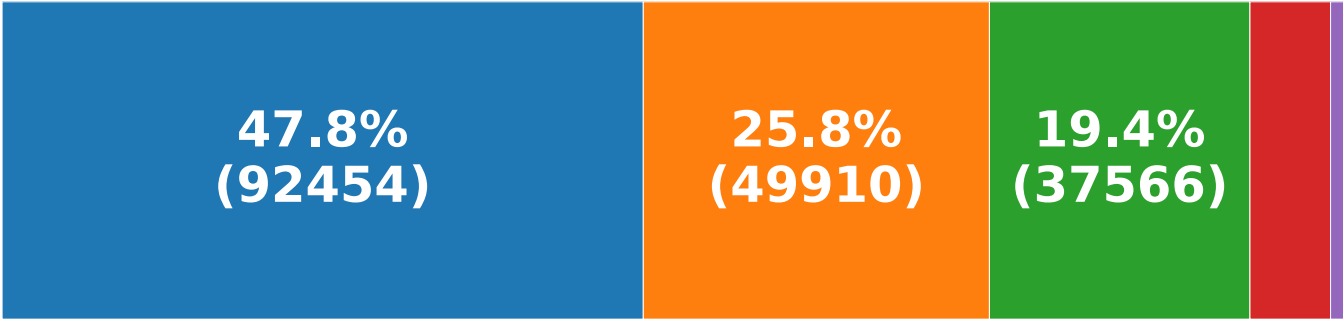
Value	Count	Frequency (%)
Ideal	92454	47.8%
Premium	49910	25.8%
Very Good	37566	19.4%
Good	11622	6.0%
Fair	2021	1.0%

Length



Histogram of lengths of the category

Common Values (Plot)



Value	Count	Frequency (%)
ideal	92454	40.0%
premium	49910	21.6%
good	49188	21.3%
very	37566	16.3%
fair	2021	0.9%

Most occurring characters

Value	Count	Frequency (%)
e	179930	14.9%
d	141642	11.8%
m	99820	8.3%
o	98376	8.2%
a	94475	7.8%
l	92454	7.7%
l	92454	7.7%
r	89497	7.4%
i	51931	4.3%
P	49910	4.1%
Other values (6)	213817	17.8%

Most occurring categories

Value	Count	Frequency (%)
Lowercase Letter	935601	77.7%
Uppercase Letter	231139	19.2%
Space Separator	37566	3.1%

Most frequent character per category

Lowercase Letter

Value	Count	Frequency (%)
e	179930	19.2%
d	141642	15.1%
m	99820	10.7%
o	98376	10.5%
a	94475	10.1%
l	92454	9.9%
r	89497	9.6%
i	51931	5.6%
u	49910	5.3%
y	37566	4.0%

Uppercase Letter

Value	Count	Frequency (%)
I	92454	40.0%
P	49910	21.6%
G	49188	21.3%
V	37566	16.3%
F	2021	0.9%

Space Separator

Value	Count	Frequency (%)
	37566	100.0%

Most occurring scripts

Value	Count	Frequency (%)
Latin	1166740	96.9%
Common	37566	3.1%

Most frequent character per script

Latin

Value	Count	Frequency (%)
e	179930	15.4%
d	141642	12.1%
m	99820	8.6%
o	98376	8.4%
a	94475	8.1%
l	92454	7.9%
l	92454	7.9%
r	89497	7.7%
i	51931	4.5%
P	49910	4.3%
Other values (5)	176251	15.1%

Common

Value	Count	Frequency (%)
	37566	100.0%

Most occurring blocks

Value	Count	Frequency (%)
ASCII	1204306	100.0%

Most frequent character per block

ASCII

Value	Count	Frequency (%)
e	179930	14.9%
d	141642	11.8%
m	99820	8.3%
o	98376	8.2%
a	94475	7.8%
l	92454	7.7%
l	92454	7.7%
r	89497	7.4%
i	51931	4.3%
P	49910	4.1%
Other values (6)	213817	17.8%

color
Categorical

Distinct	7
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	10.7 MiB

Length

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

Characters and Unicode

Total characters	193573	
Distinct characters	7	
Distinct categories	1 (https://en.wikipedia.org/wiki/Unicode_character_property#General_Category)	?
Distinct scripts	1 (https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode)	?
Distinct blocks	1 (https://en.wikipedia.org/wiki/Unicode_block)	?

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

Unique

Unique	0	?
Unique (%)	0.0%	

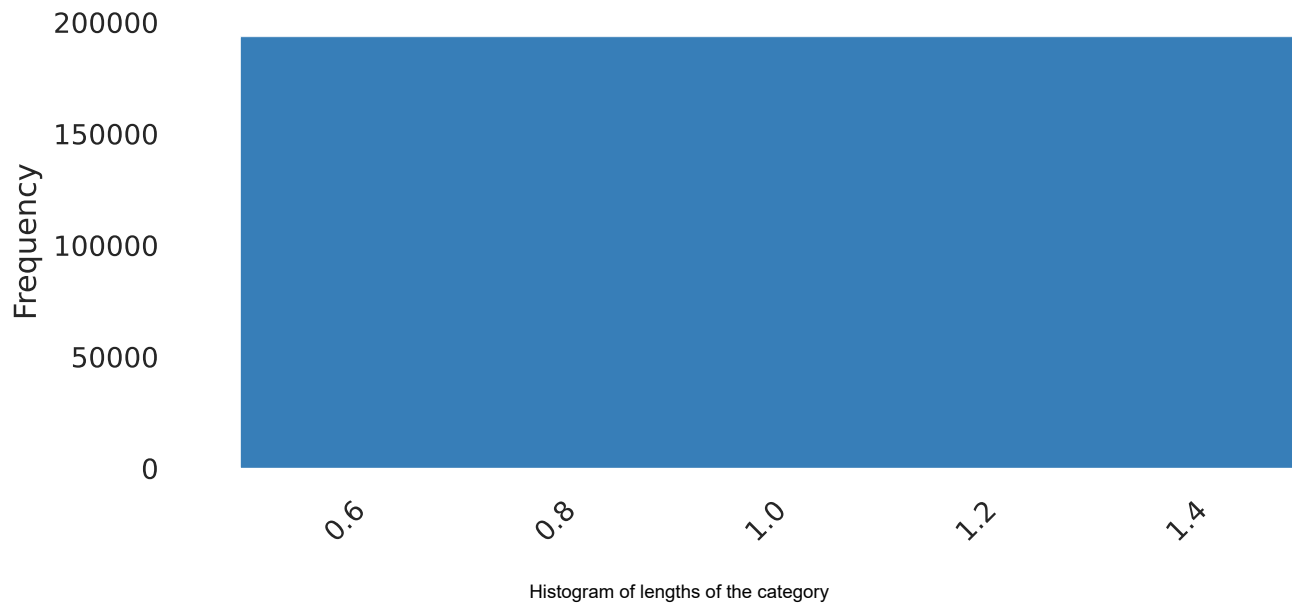
Sample

1st row	F
2nd row	J
3rd row	G
4th row	G
5th row	G

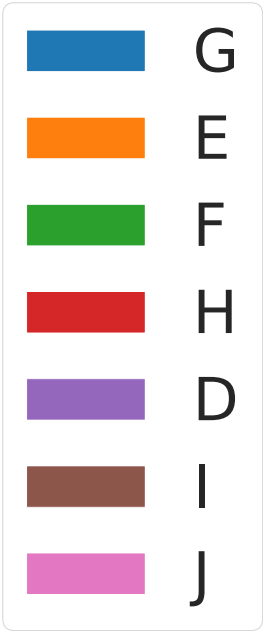
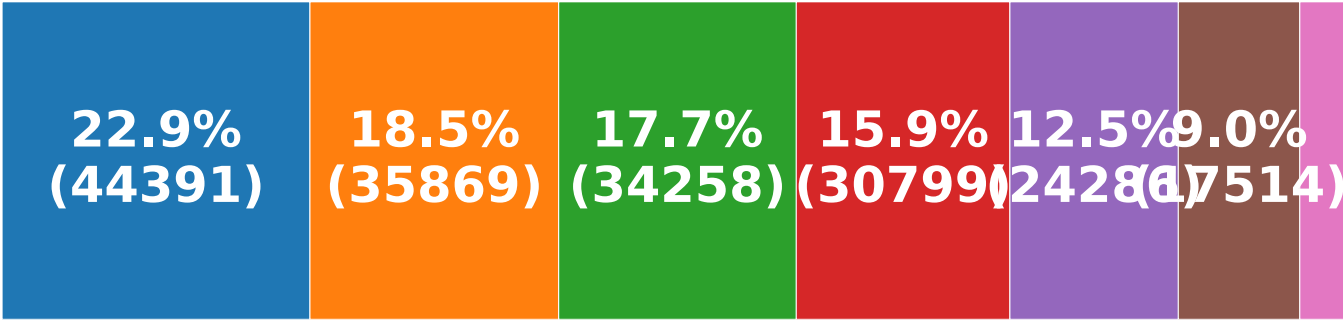
Common Values

Value	Count	Frequency (%)
G	44391	22.9%
E	35869	18.5%
F	34258	17.7%
H	30799	15.9%
D	24286	12.5%
I	17514	9.0%
J	6456	3.3%

Length



Common Values (Plot)



Value	Count	Frequency (%)
g	44391	22.9%
e	35869	18.5%
f	34258	17.7%
h	30799	15.9%
d	24286	12.5%
i	17514	9.0%
j	6456	3.3%

Most occurring characters

Value	Count	Frequency (%)
G	44391	22.9%
E	35869	18.5%
F	34258	17.7%
H	30799	15.9%
D	24286	12.5%
I	17514	9.0%
J	6456	3.3%

Most occurring categories

Value	Count	Frequency (%)
Uppercase Letter	193573	100.0%

Most frequent character per category

Uppercase Letter

Value	Count	Frequency (%)
G	44391	22.9%
E	35869	18.5%
F	34258	17.7%
H	30799	15.9%
D	24286	12.5%
I	17514	9.0%
J	6456	3.3%

Most occurring scripts

Value	Count	Frequency (%)
Latin	193573	100.0%

Most frequent character per script

Latin

Value	Count	Frequency (%)
G	44391	22.9%
E	35869	18.5%
F	34258	17.7%
H	30799	15.9%
D	24286	12.5%
I	17514	9.0%
J	6456	3.3%

Most occurring blocks

Value	Count	Frequency (%)
ASCII	193573	100.0%

Most frequent character per block

ASCII

Value	Count	Frequency (%)
G	44391	22.9%
E	35869	18.5%
F	34258	17.7%
H	30799	15.9%
D	24286	12.5%
I	17514	9.0%
J	6456	3.3%

clarity
Categorical

Distinct	8
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	11.1 MiB

Length

Max length	4
Median length	3
Mean length	3.1118906
Min length	2

Characters and Unicode

Total characters	602378	
Distinct characters	6	
Distinct categories	2 (https://en.wikipedia.org/wiki/Unicode_character_property#General_Category)	?
Distinct scripts	2 (https://en.wikipedia.org/wiki/Script_(Unicode)#List_of_scripts_in_Unicode)	?
Distinct blocks	1 (https://en.wikipedia.org/wiki/Unicode_block)	?

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

Unique

Unique	0	?
Unique (%)	0.0%	

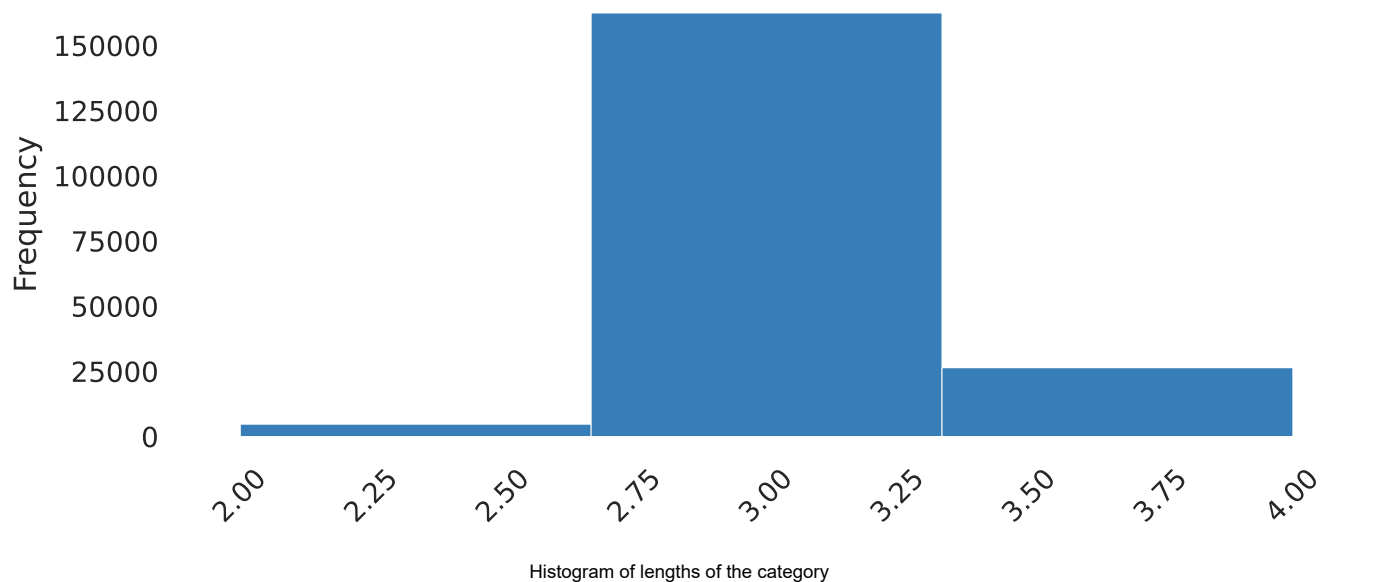
Sample

1st row	VS2
2nd row	SI2
3rd row	VS1
4th row	VS1
5th row	VS2

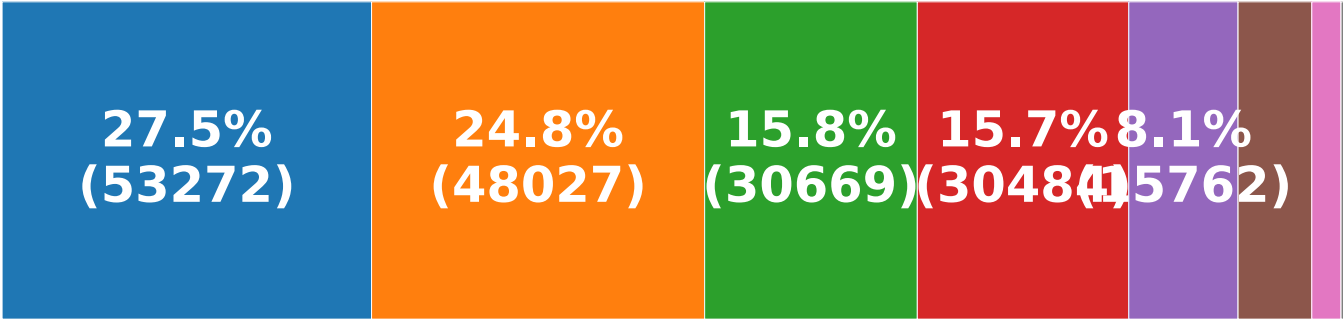
Common Values

Value	Count	Frequency (%)
SI1	53272	27.5%
VS2	48027	24.8%
VS1	30669	15.8%
SI2	30484	15.7%
VVS2	15762	8.1%
VVS1	10628	5.5%
IF	4219	2.2%
I1	512	0.3%

Length



Common Values (Plot)



Value	Count	Frequency (%)
si1	53272	27.5%
vs2	48027	24.8%
vs1	30669	15.8%
si2	30484	15.7%
vvs2	15762	8.1%
vvs1	10628	5.5%
if	4219	2.2%
i1	512	0.3%

Most occurring characters

Value	Count	Frequency (%)
S	188842	31.3%
V	131476	21.8%
1	95081	15.8%
2	94273	15.7%
I	88487	14.7%
F	4219	0.7%

Most occurring categories

Value	Count	Frequency (%)
Uppercase Letter	413024	68.6%
Decimal Number	189354	31.4%

Most frequent character per category

Uppercase Letter

Value	Count	Frequency (%)
S	188842	45.7%
V	131476	31.8%
I	88487	21.4%
F	4219	1.0%

Decimal Number

Value	Count	Frequency (%)
1	95081	50.2%
2	94273	49.8%

Most occurring scripts

Value	Count	Frequency (%)
Latin	413024	68.6%
Common	189354	31.4%

Most frequent character per script

Latin

Value	Count	Frequency (%)
S	188842	45.7%
V	131476	31.8%
I	88487	21.4%
F	4219	1.0%

Common

Value	Count	Frequency (%)
1	95081	50.2%
2	94273	49.8%

Most occurring blocks

Value	Count	Frequency (%)
ASCII	602378	100.0%

Most frequent character per block

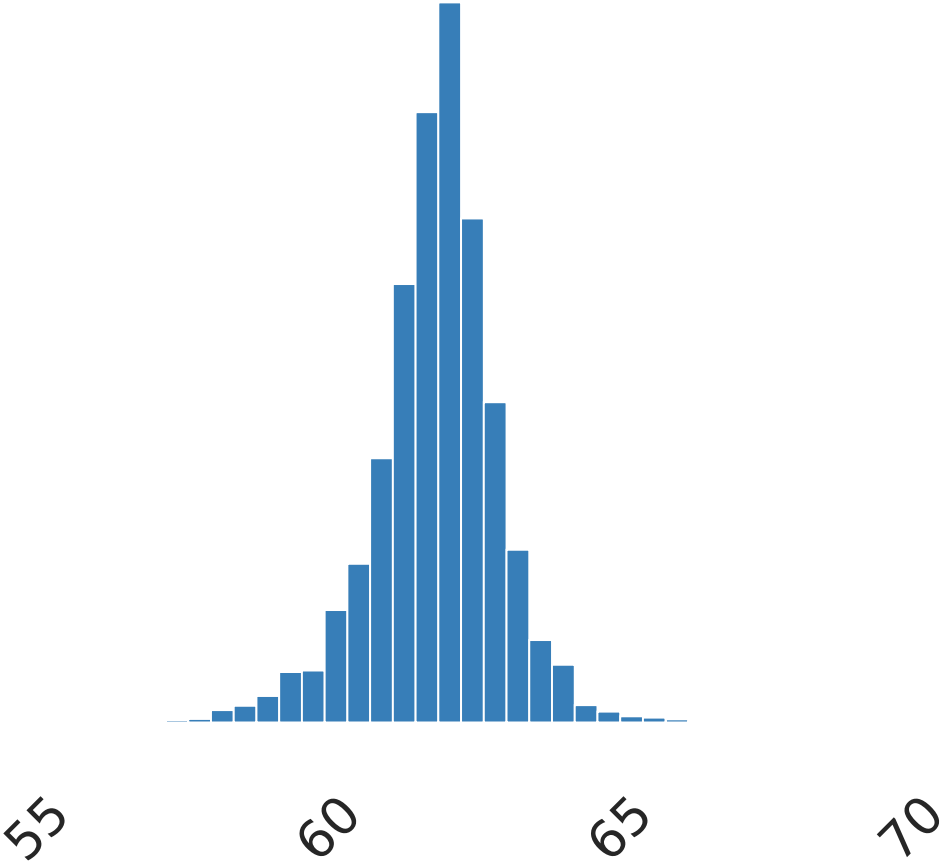
ASCII

Value	Count	Frequency (%)
S	188842	31.3%
V	131476	21.8%
1	95081	15.8%
2	94273	15.7%
I	88487	14.7%
F	4219	0.7%

depth

Real number (\mathbb{R})

Distinct	153
Distinct (%)	0.1%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Mean	61.820574
Minimum	52.1
Maximum	71.6
Zeros	0
Zeros (%)	0.0%
Negative	0
Negative (%)	0.0%
Memory size	1.5 MiB

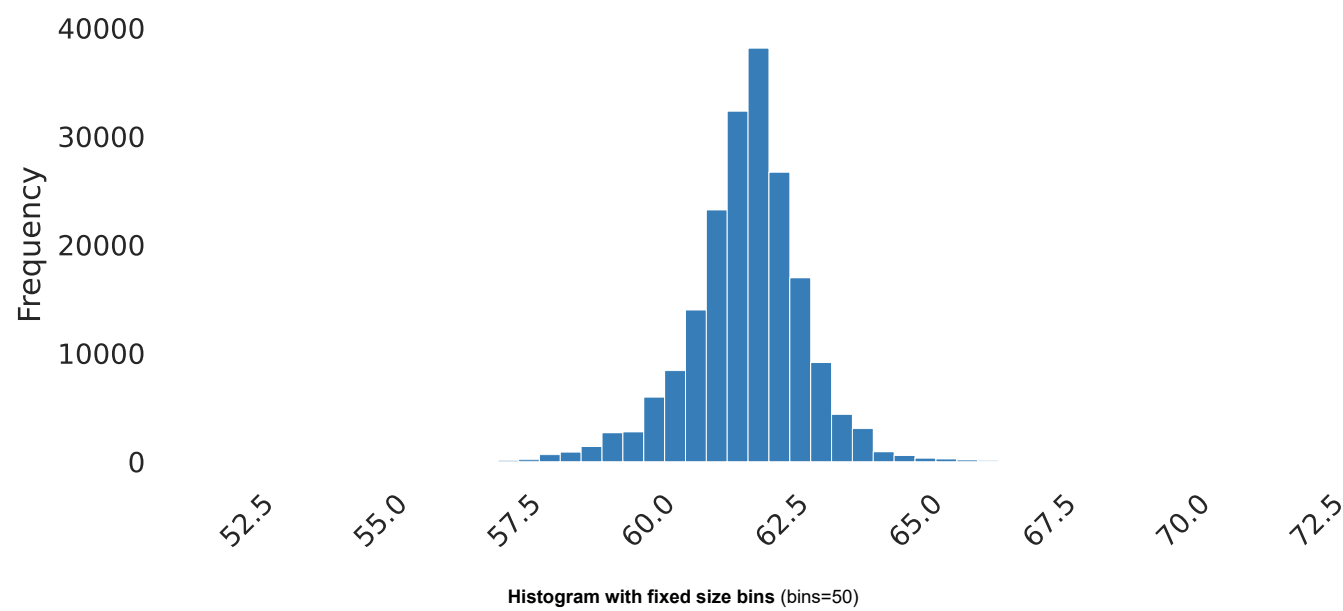


Quantile statistics

Minimum	52.1
5-th percentile	59.9
Q1	61.3
median	61.9
Q3	62.4
95-th percentile	63.5
Maximum	71.6
Range	19.5
Interquartile range (IQR)	1.1

Descriptive statistics

Standard deviation	1.0817044
Coefficient of variation (CV)	0.017497482
Kurtosis	2.4770411
Mean	61.820574
Median Absolute Deviation (MAD)	0.6
Skewness	-0.27638236
Sum	11966794
Variance	1.1700843
Monotonicity	Not monotonic



Value	Count	Frequency (%)
61.9	10781	5.6%
62	10150	5.2%
61.8	9270	4.8%
62.1	8866	4.6%
61.6	8534	4.4%
62.2	8345	4.3%
62.3	7987	4.1%
61.7	7970	4.1%
62.4	7030	3.6%
61.5	6554	3.4%
Other values (143)	108086	55.8%

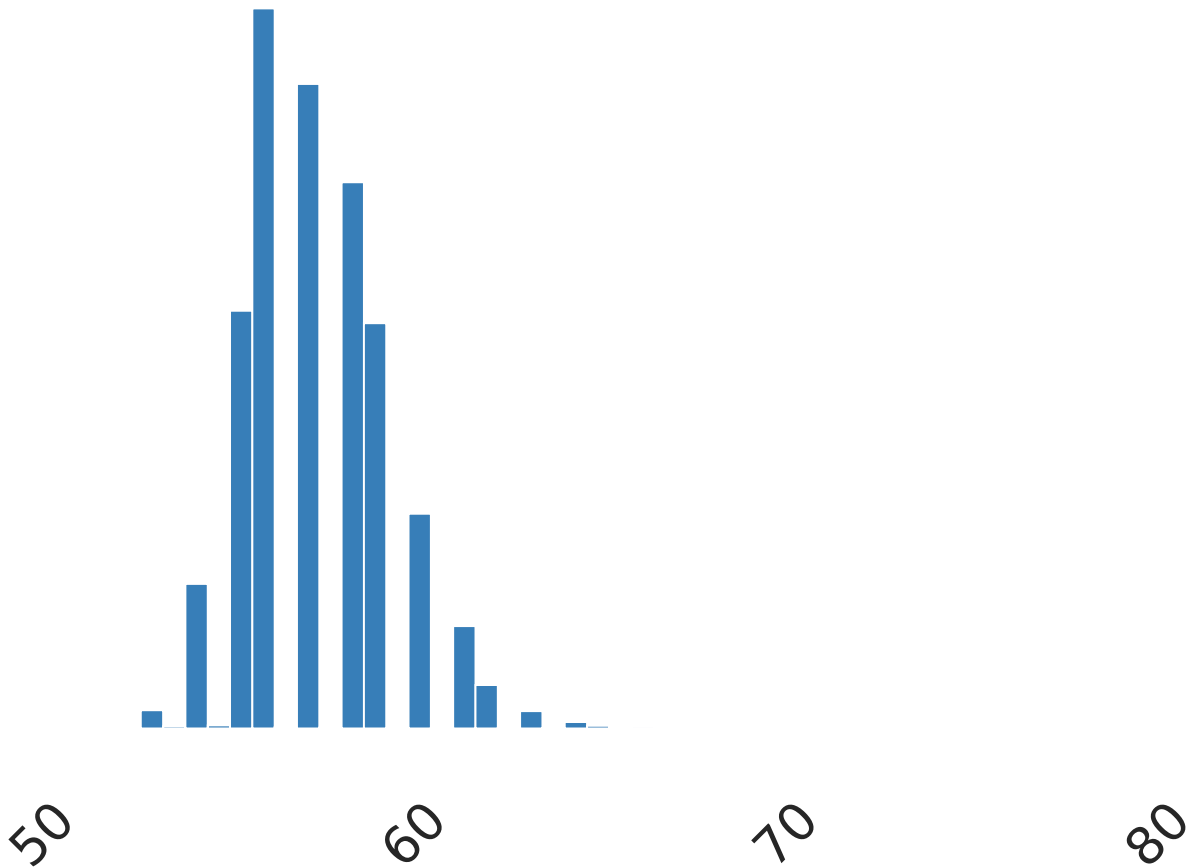
Value	Count	Frequency (%)
52.1	1	< 0.1%
52.2	1	< 0.1%
52.7	1	< 0.1%
53.1	2	< 0.1%
53.2	2	< 0.1%
54.7	1	< 0.1%
54.9	2	< 0.1%
55	1	< 0.1%
55.1	1	< 0.1%
55.2	5	< 0.1%

Value	Count	Frequency (%)
71.6	2	< 0.1%
70	1	< 0.1%
69.9	1	< 0.1%
69.6	1	< 0.1%
69.5	3	< 0.1%
69.4	1	< 0.1%
69.2	2	< 0.1%
69.1	1	< 0.1%
69	1	< 0.1%
68.9	1	< 0.1%

table

Real number (R)

Distinct	108
Distinct (%)	0.1%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Mean	57.227675
Minimum	49
Maximum	79
Zeros	0
Zeros (%)	0.0%
Negative	0
Negative (%)	0.0%
Memory size	1.5 MiB

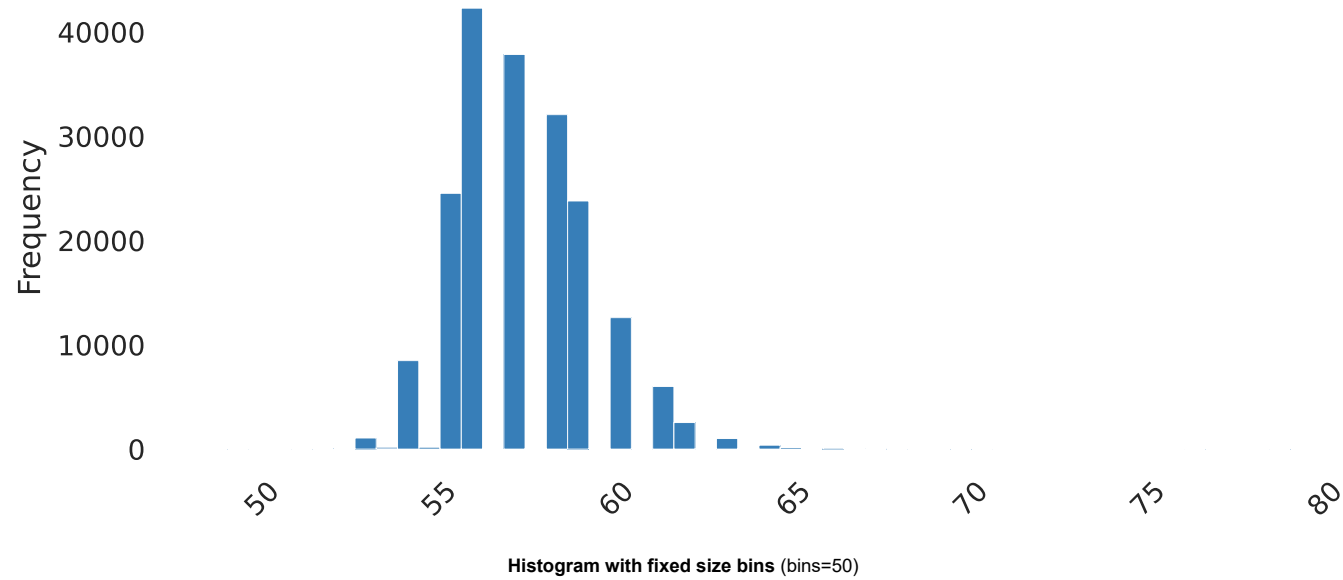


Quantile statistics

Minimum	49
5-th percentile	54.2
Q1	56
median	57
Q3	58
95-th percentile	61
Maximum	79
Range	30
Interquartile range (IQR)	2

Descriptive statistics

Standard deviation	1.9188443
Coefficient of variation (CV)	0.033530006
Kurtosis	0.81018001
Mean	57.227675
Median Absolute Deviation (MAD)	1
Skewness	0.61906223
Sum	11077733
Variance	3.6819634
Monotonicity	Not monotonic



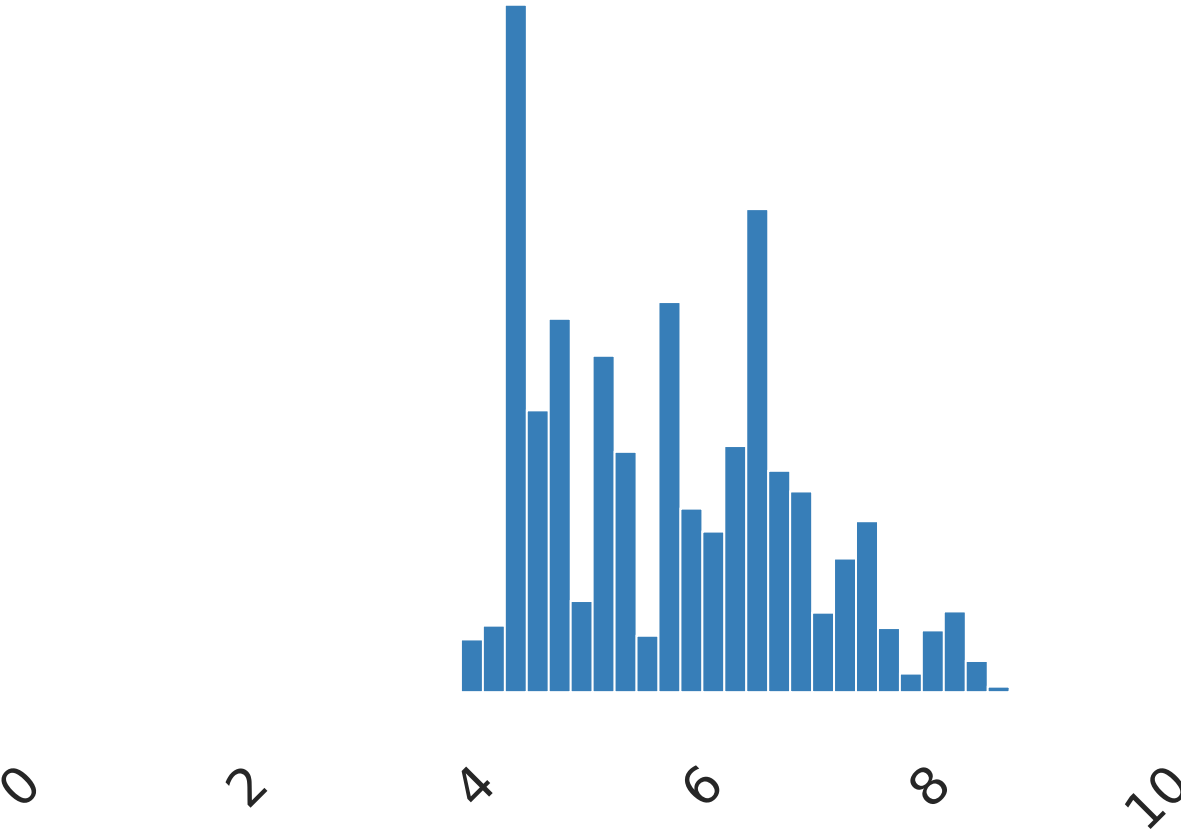
Value	Count	Frequency (%)
56	42194	21.8%
57	37827	19.5%
58	32045	16.6%
55	24429	12.6%
59	23784	12.3%
60	12584	6.5%
54	8281	4.3%
61	6002	3.1%
62	2545	1.3%
53	1069	0.6%
Other values (98)	2813	1.5%

Value	Count	Frequency (%)
49	1	< 0.1%
51	6	< 0.1%
52	50	< 0.1%
53	1069	0.6%
53.1	1	< 0.1%
53.2	8	< 0.1%
53.3	17	< 0.1%
53.4	7	< 0.1%
53.5	10	< 0.1%
53.6	40	< 0.1%

Value	Count	Frequency (%)
79	1	< 0.1%
76	1	< 0.1%
70	5	< 0.1%
69	10	< 0.1%
68	20	< 0.1%
67	32	< 0.1%
66	114	0.1%
65	157	0.1%
64.4	1	< 0.1%
64	376	0.2%

X
Real number (\mathbb{R})

Distinct	522
Distinct (%)	0.3%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Mean	5.7153121
Minimum	0
Maximum	9.65
Zeros	3
Zeros (%)	< 0.1%
Negative	0
Negative (%)	0.0%
Memory size	1.5 MiB

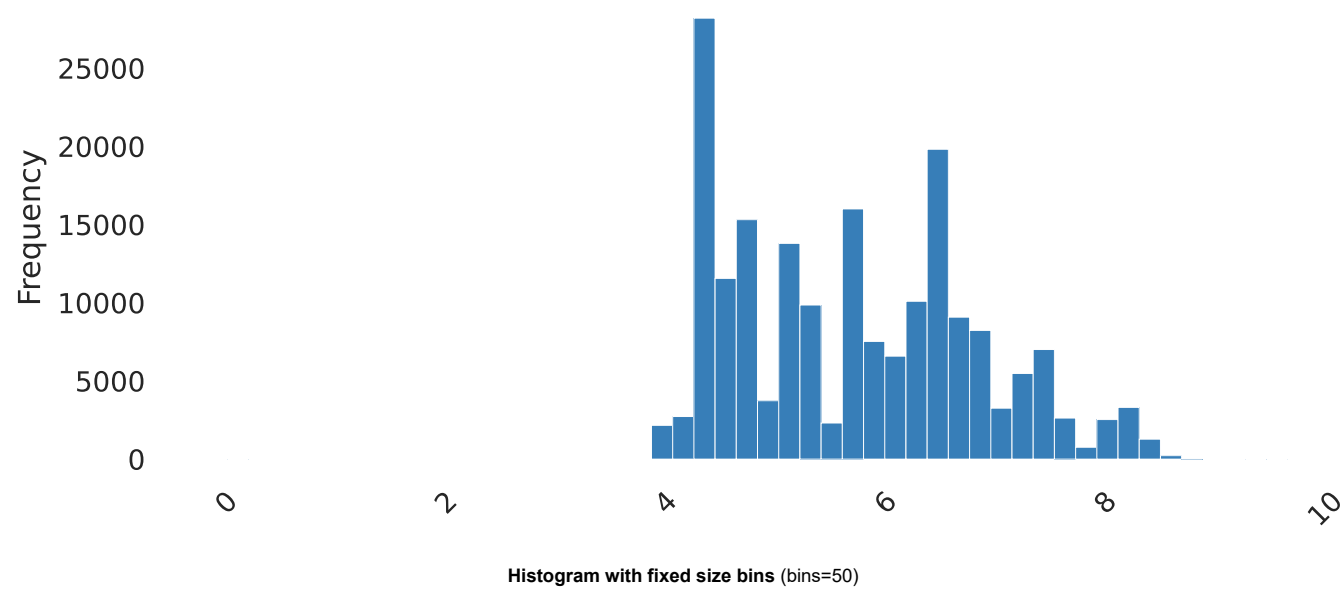


Quantile statistics

Minimum	0
5-th percentile	4.29
Q1	4.7
median	5.7
Q3	6.51
95-th percentile	7.58
Maximum	9.65
Range	9.65
Interquartile range (IQR)	1.81

Descriptive statistics

Standard deviation	1.1094222
Coefficient of variation (CV)	0.19411401
Kurtosis	-0.80100603
Mean	5.7153121
Median Absolute Deviation (MAD)	0.92
Skewness	0.36104978
Sum	1106330.1
Variance	1.2308175
Monotonicity	Not monotonic



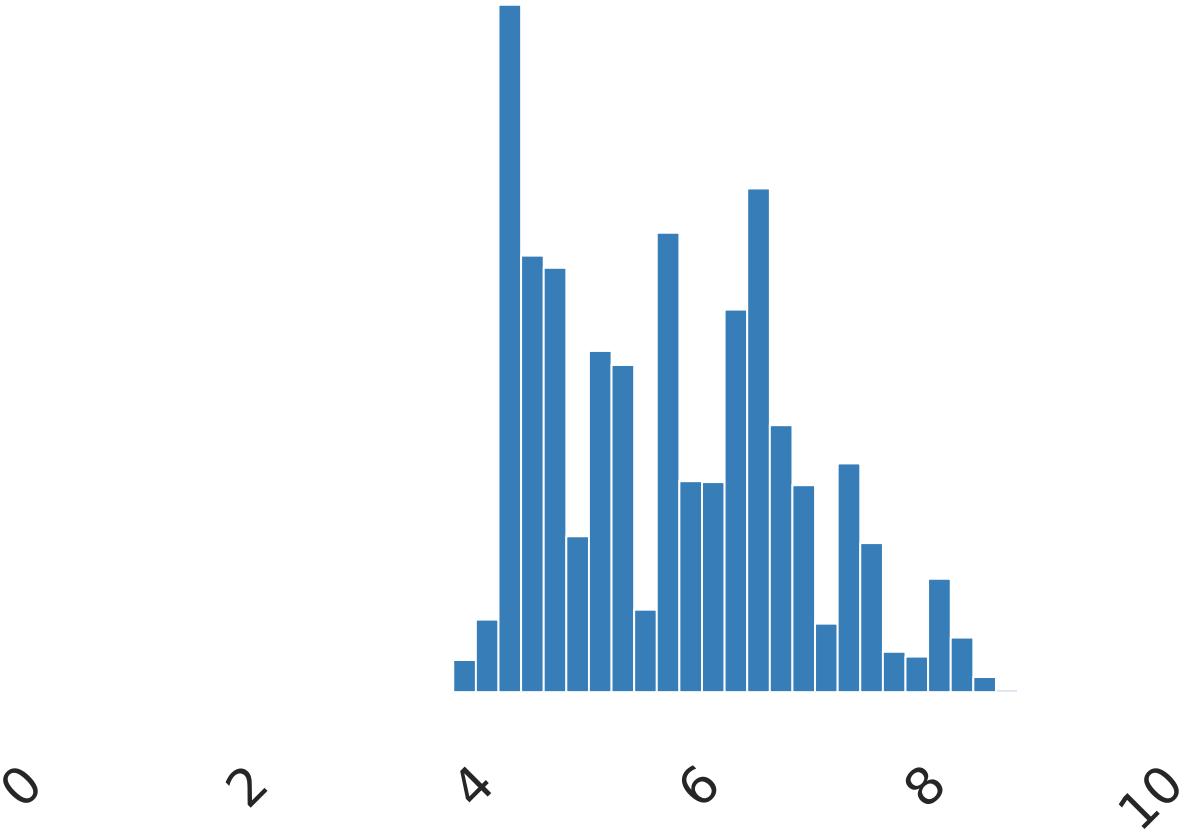
Value	Count	Frequency (%)
4.32	2094	1.1%
4.34	2010	1.0%
4.38	1986	1.0%
4.37	1938	1.0%
4.33	1812	0.9%
4.31	1761	0.9%
4.35	1685	0.9%
4.41	1644	0.8%
4.36	1568	0.8%
4.3	1558	0.8%
Other values (512)	175517	90.7%

Value	Count	Frequency (%)
0	3	< 0.1%
3.75	2	< 0.1%
3.77	7	< 0.1%
3.78	5	< 0.1%
3.79	4	< 0.1%
3.81	7	< 0.1%
3.82	6	< 0.1%
3.84	9	< 0.1%
3.85	2	< 0.1%
3.86	9	< 0.1%

Value	Count	Frequency (%)
9.65	1	< 0.1%
9.51	1	< 0.1%
9.46	1	< 0.1%
9.43	1	< 0.1%
9.42	4	< 0.1%
9.36	3	< 0.1%
9.35	1	< 0.1%
9.3	1	< 0.1%
9.24	1	< 0.1%
9.1	1	< 0.1%

y
Real number (\mathbb{R})

Distinct	521
Distinct (%)	0.3%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Mean	5.7200944
Minimum	0
Maximum	10.01
Zeros	2
Zeros (%)	< 0.1%
Negative	0
Negative (%)	0.0%
Memory size	1.5 MiB

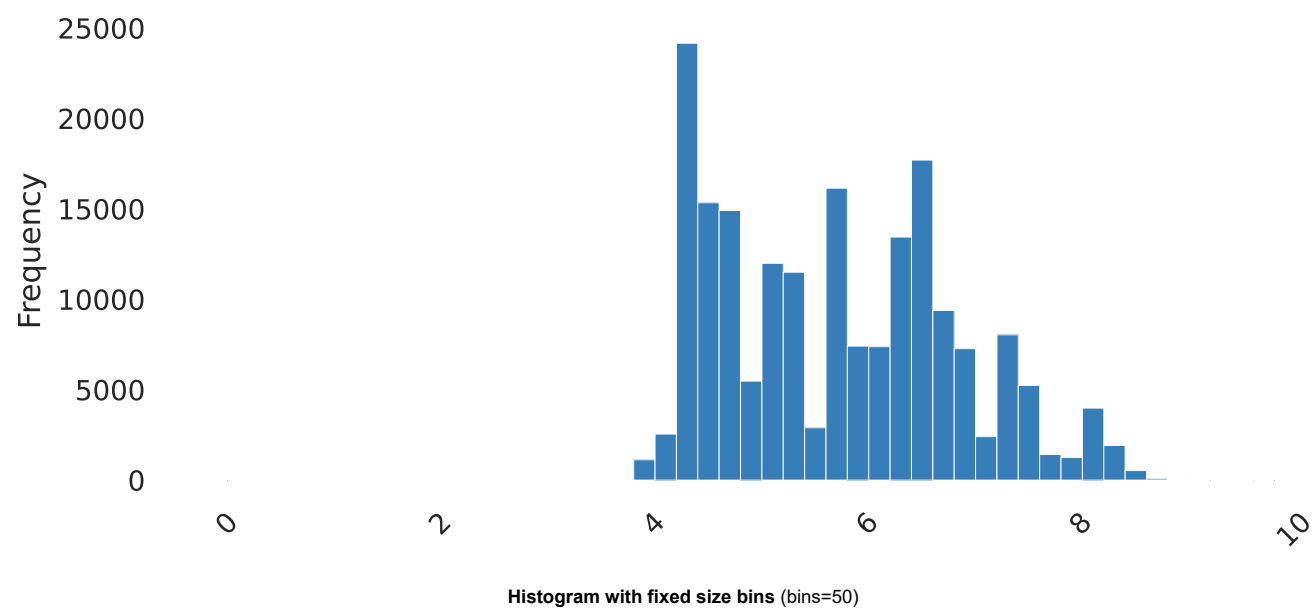


Quantile statistics

Minimum	0
5-th percentile	4.31
Q1	4.71
median	5.72
Q3	6.51
95-th percentile	7.58
Maximum	10.01
Range	10.01
Interquartile range (IQR)	1.8

Descriptive statistics

Standard deviation	1.1023335
Coefficient of variation (CV)	0.19271246
Kurtosis	-0.81066804
Mean	5.7200944
Median Absolute Deviation (MAD)	0.92
Skewness	0.35675813
Sum	1107255.8
Variance	1.2151391
Monotonicity	Not monotonic



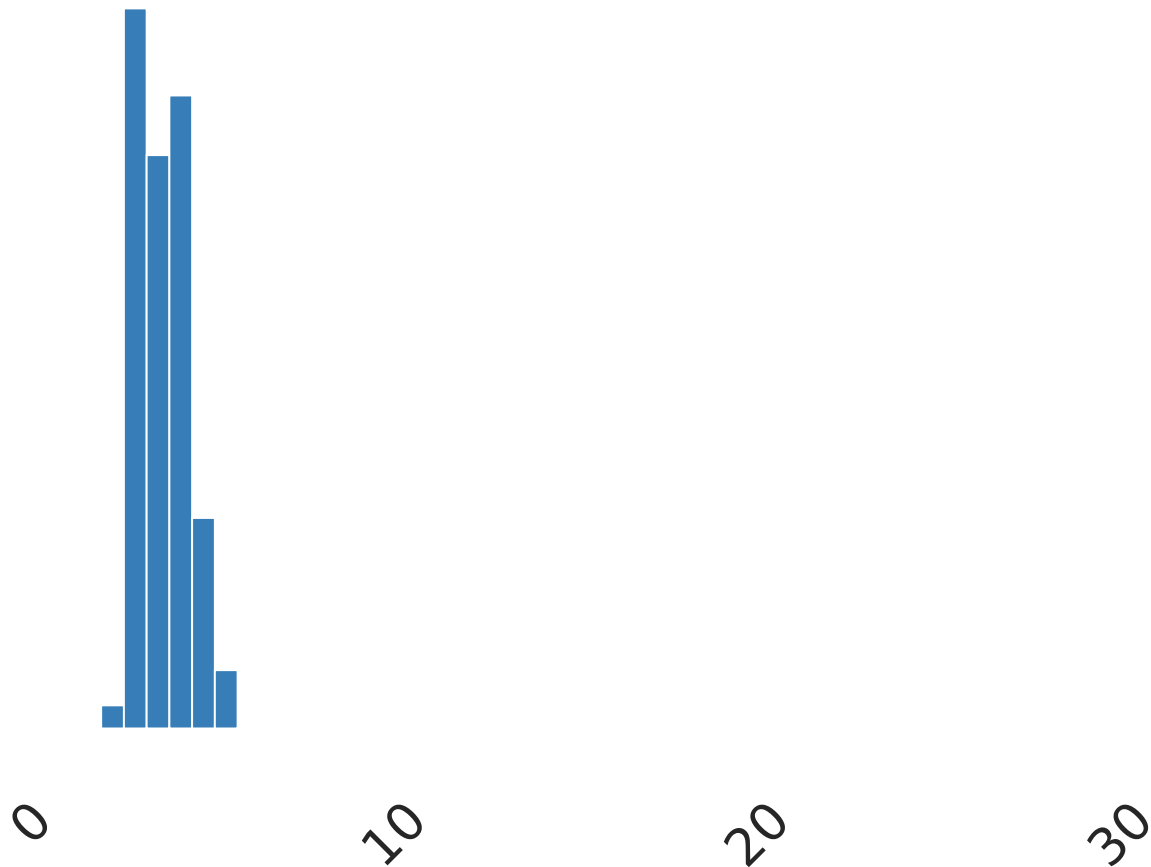
Value	Count	Frequency (%)
4.38	2163	1.1%
4.35	2088	1.1%
4.34	1986	1.0%
4.37	1911	1.0%
4.31	1854	1.0%
4.39	1719	0.9%
4.33	1711	0.9%
4.36	1625	0.8%
4.32	1618	0.8%
4.41	1593	0.8%
Other values (511)	175305	90.6%

Value	Count	Frequency (%)
0	2	< 0.1%
3.71	1	< 0.1%
3.72	7	< 0.1%
3.73	1	< 0.1%
3.75	2	< 0.1%
3.77	4	< 0.1%
3.78	9	< 0.1%
3.79	2	< 0.1%
3.8	1	< 0.1%
3.81	5	< 0.1%

Value	Count	Frequency (%)
10.01	1	< 0.1%
9.59	1	< 0.1%
9.46	1	< 0.1%
9.36	1	< 0.1%
9.34	1	< 0.1%
9.31	1	< 0.1%
9.3	2	< 0.1%
9.26	2	< 0.1%
9.24	2	< 0.1%
9.14	1	< 0.1%

Z
Real number (\mathbb{R})

Distinct	349
Distinct (%)	0.2%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Mean	3.5342463
Minimum	0
Maximum	31.3
Zeros	10
Zeros (%)	< 0.1%
Negative	0
Negative (%)	0.0%
Memory size	1.5 MiB

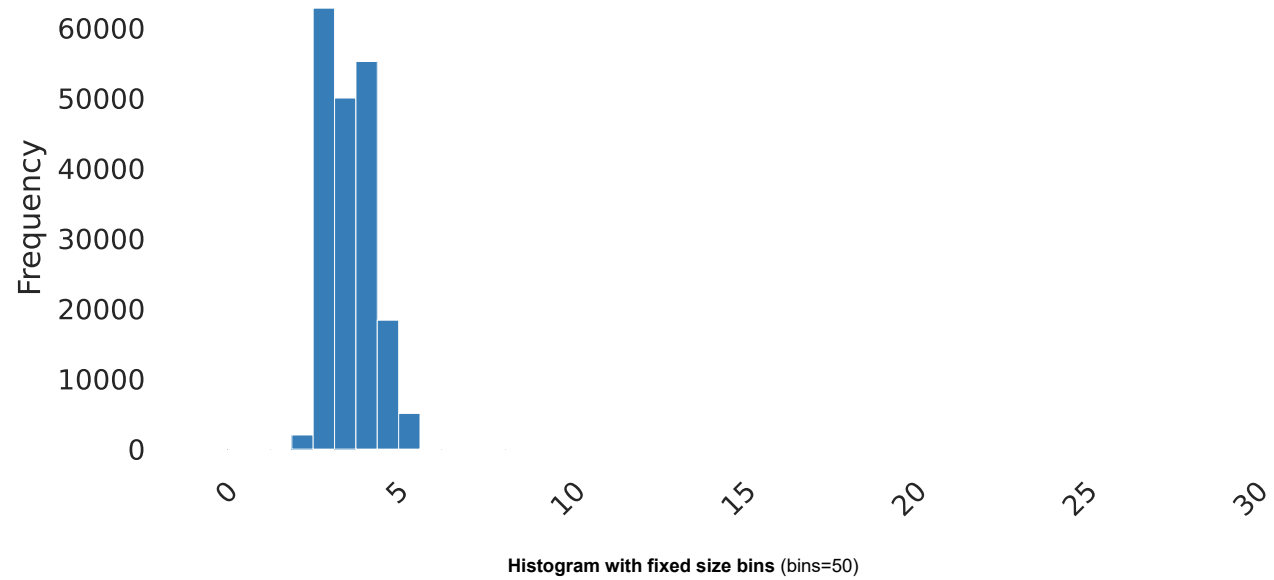


Quantile statistics

Minimum	0
5-th percentile	2.66
Q1	2.9
median	3.53
Q3	4.03
95-th percentile	4.684
Maximum	31.3
Range	31.3
Interquartile range (IQR)	1.13

Descriptive statistics

Standard deviation	0.68892211
Coefficient of variation (CV)	0.19492759
Kurtosis	12.818313
Mean	3.5342463
Median Absolute Deviation (MAD)	0.57
Skewness	0.68567148
Sum	684134.67
Variance	0.47461367
Monotonicity	Not monotonic



Value	Count	Frequency (%)
2.69	3523	1.8%
2.7	3305	1.7%
2.68	3289	1.7%
2.71	3087	1.6%
2.72	2887	1.5%
2.73	2747	1.4%
2.67	2628	1.4%
3.99	2542	1.3%
4.02	2430	1.3%
4.01	2384	1.2%
Other values (339)	164751	85.1%

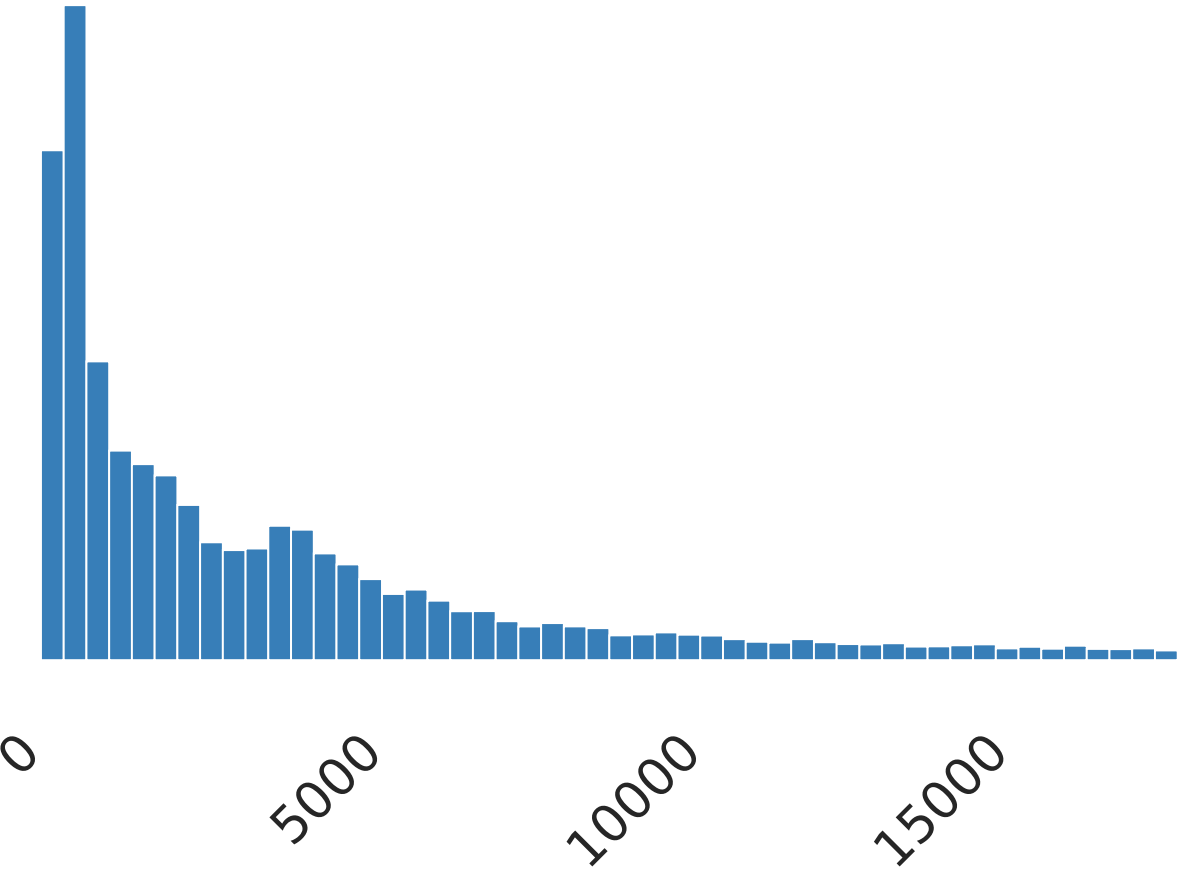
Value	Count	Frequency (%)
0	10	< 0.1%
1.05	1	< 0.1%
2.24	2	< 0.1%
2.26	1	< 0.1%
2.27	2	< 0.1%
2.28	1	< 0.1%
2.3	4	< 0.1%
2.31	9	< 0.1%
2.32	5	< 0.1%
2.33	7	< 0.1%

Value	Count	Frequency (%)
31.3	1	< 0.1%
8.4	1	< 0.1%
8.35	1	< 0.1%
8.18	1	< 0.1%
6.03	1	< 0.1%
5.75	1	< 0.1%
5.73	1	< 0.1%
5.69	1	< 0.1%
5.67	2	< 0.1%
5.65	1	< 0.1%

price

Real number (R)

Distinct	8738
Distinct (%)	4.5%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Mean	3969.1554
Minimum	326
Maximum	18818
Zeros	0
Zeros (%)	0.0%
Negative	0
Negative (%)	0.0%
Memory size	1.5 MiB

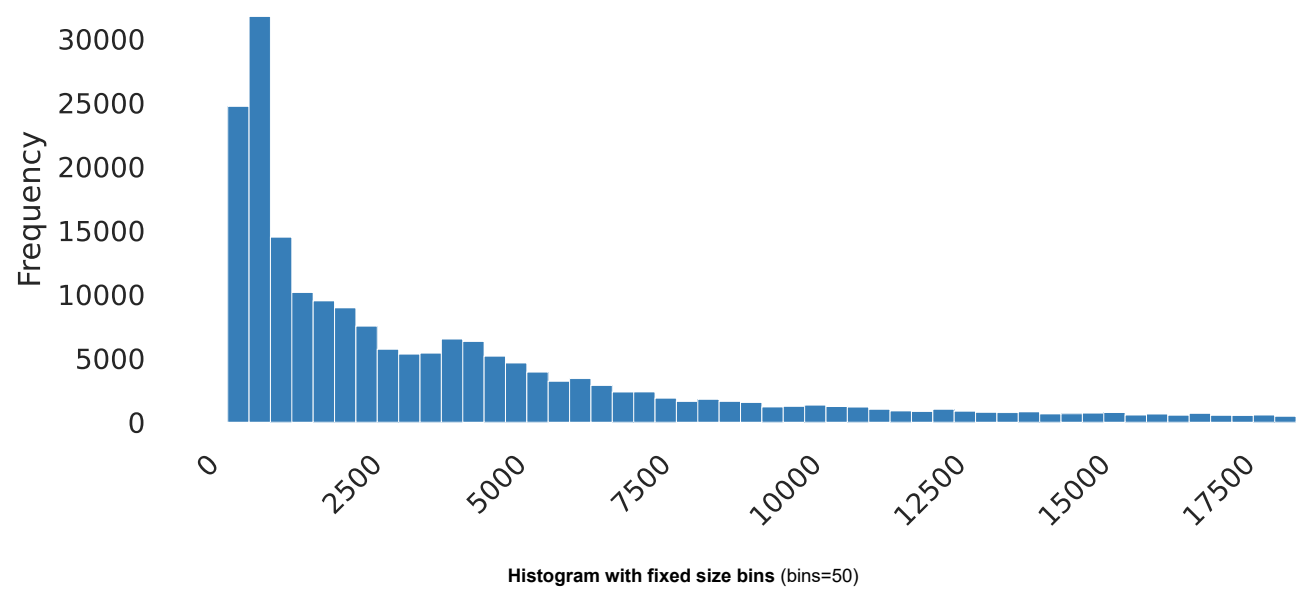


Quantile statistics

Minimum	326
5-th percentile	544
Q1	951
median	2401
Q3	5408
95-th percentile	13298
Maximum	18818
Range	18492
Interquartile range (IQR)	4457

Descriptive statistics

Standard deviation	4034.3741
Coefficient of variation (CV)	1.0164314
Kurtosis	2.106914
Mean	3969.1554
Median Absolute Deviation (MAD)	1678
Skewness	1.6055812
Sum	7.6832132×10^8
Variance	16276175
Monotonicity	Not monotonic

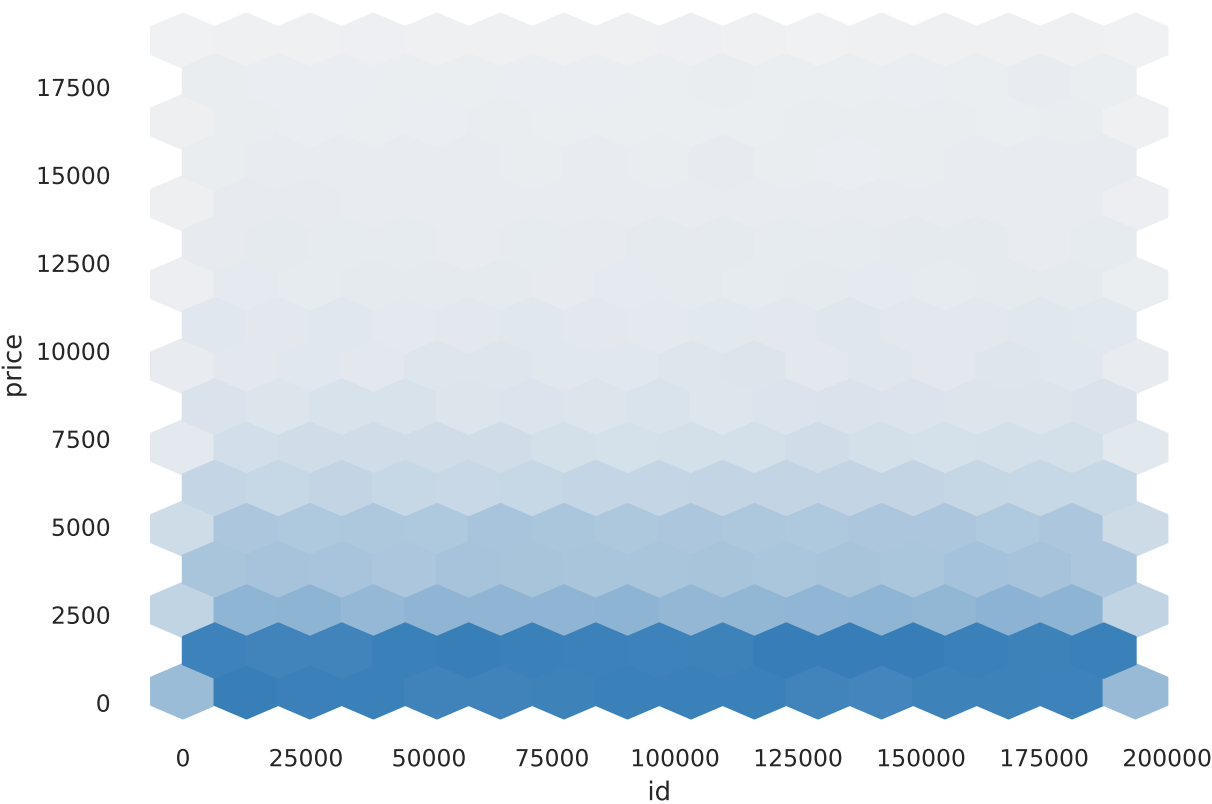


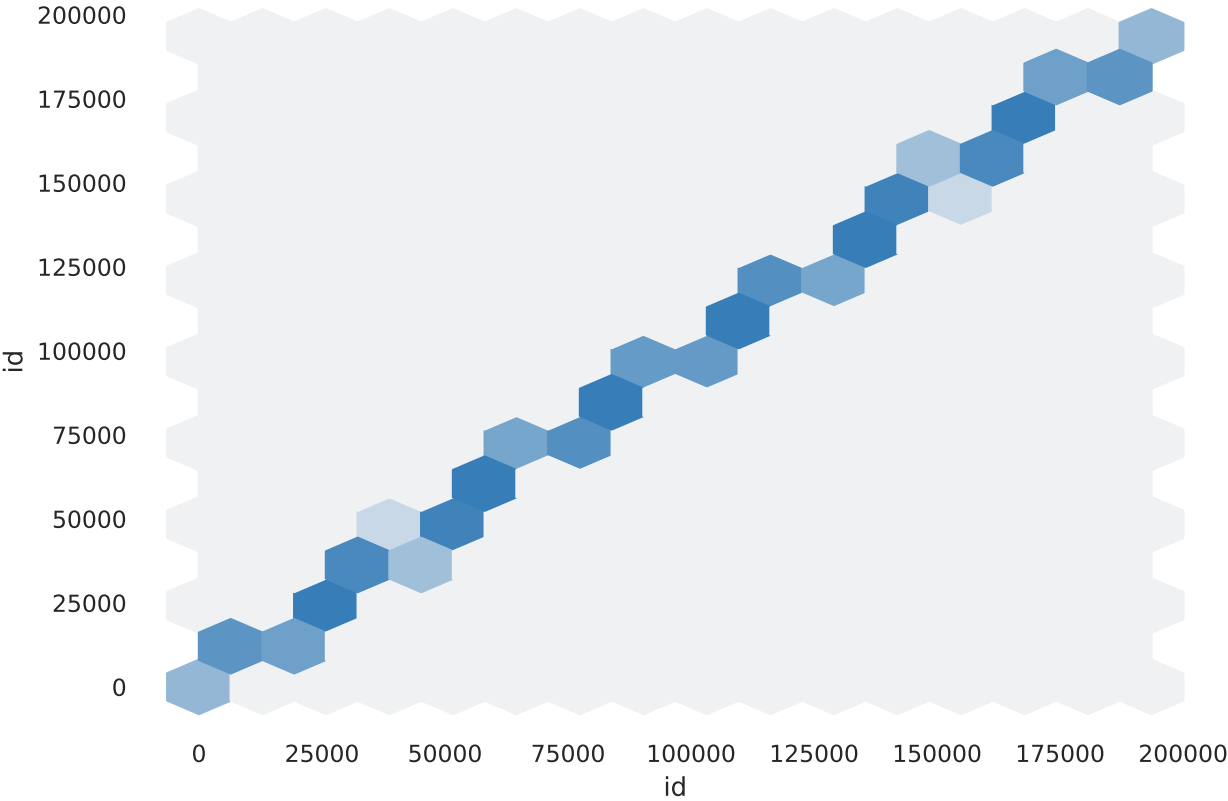
Value	Count	Frequency (%)
544	542	0.3%
605	464	0.2%
789	454	0.2%
828	438	0.2%
776	437	0.2%
802	435	0.2%
552	427	0.2%
561	416	0.2%
625	414	0.2%
596	401	0.2%
Other values (8728)	189145	97.7%

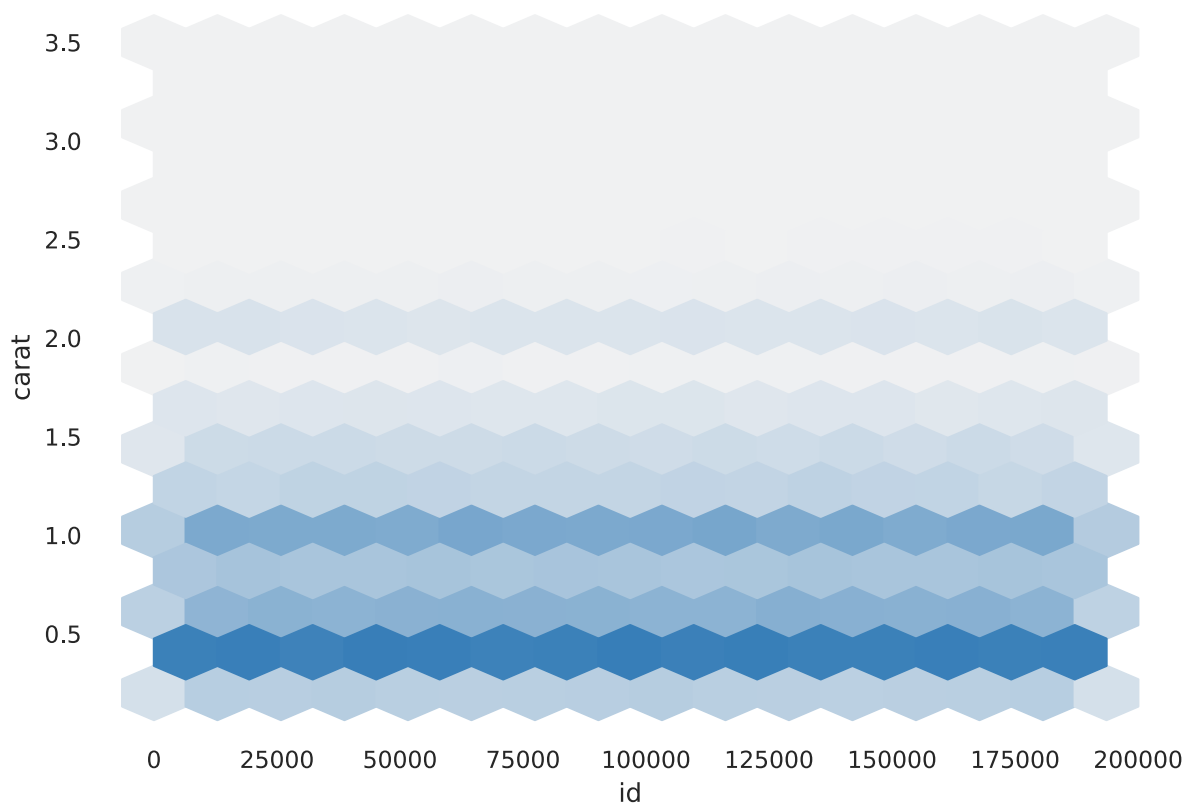
Value	Count	Frequency (%)
326	15	< 0.1%
335	11	< 0.1%
336	11	< 0.1%
337	5	< 0.1%
338	11	< 0.1%
344	5	< 0.1%
345	15	< 0.1%
348	5	< 0.1%
351	25	< 0.1%
357	31	< 0.1%

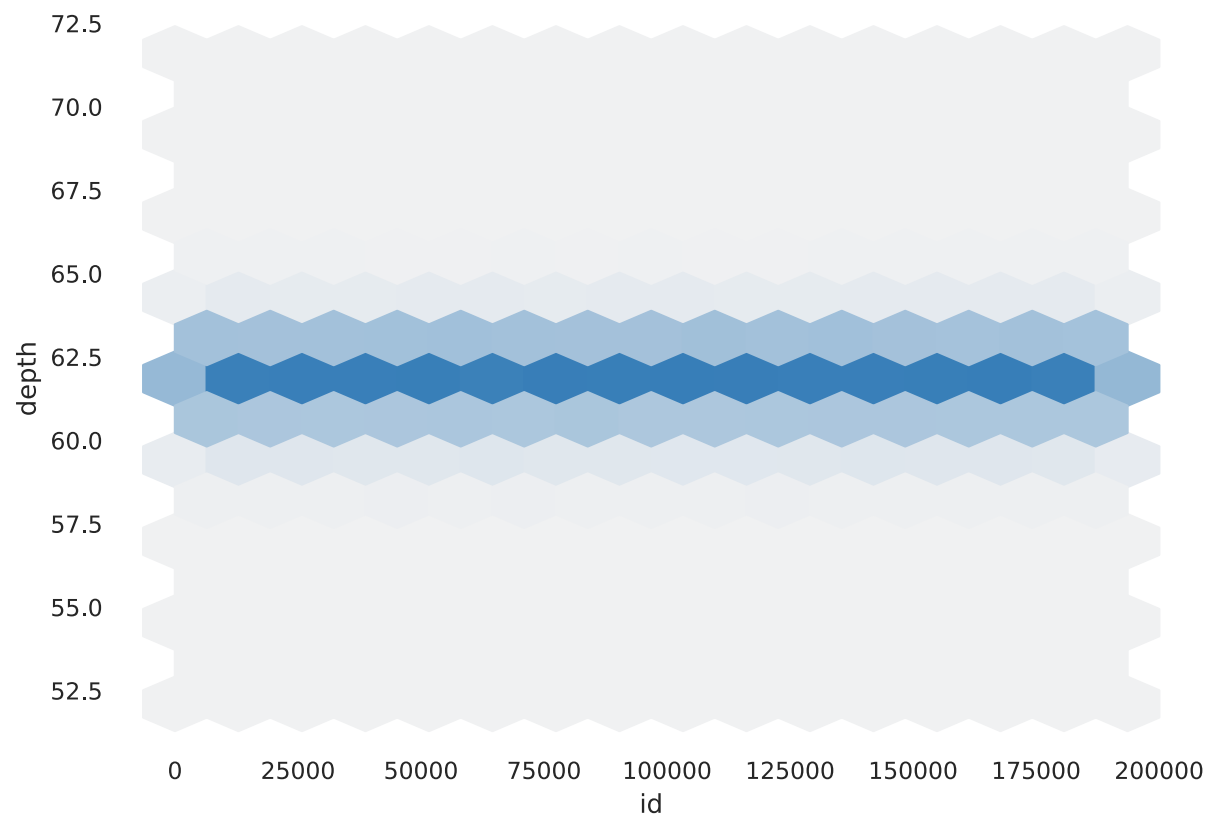
Value	Count	Frequency (%)
18818	6	< 0.1%
18804	9	< 0.1%
18795	10	< 0.1%
18791	15	< 0.1%
18787	5	< 0.1%
18780	6	< 0.1%
18766	12	< 0.1%
18760	6	< 0.1%
18759	9	< 0.1%
18757	8	< 0.1%

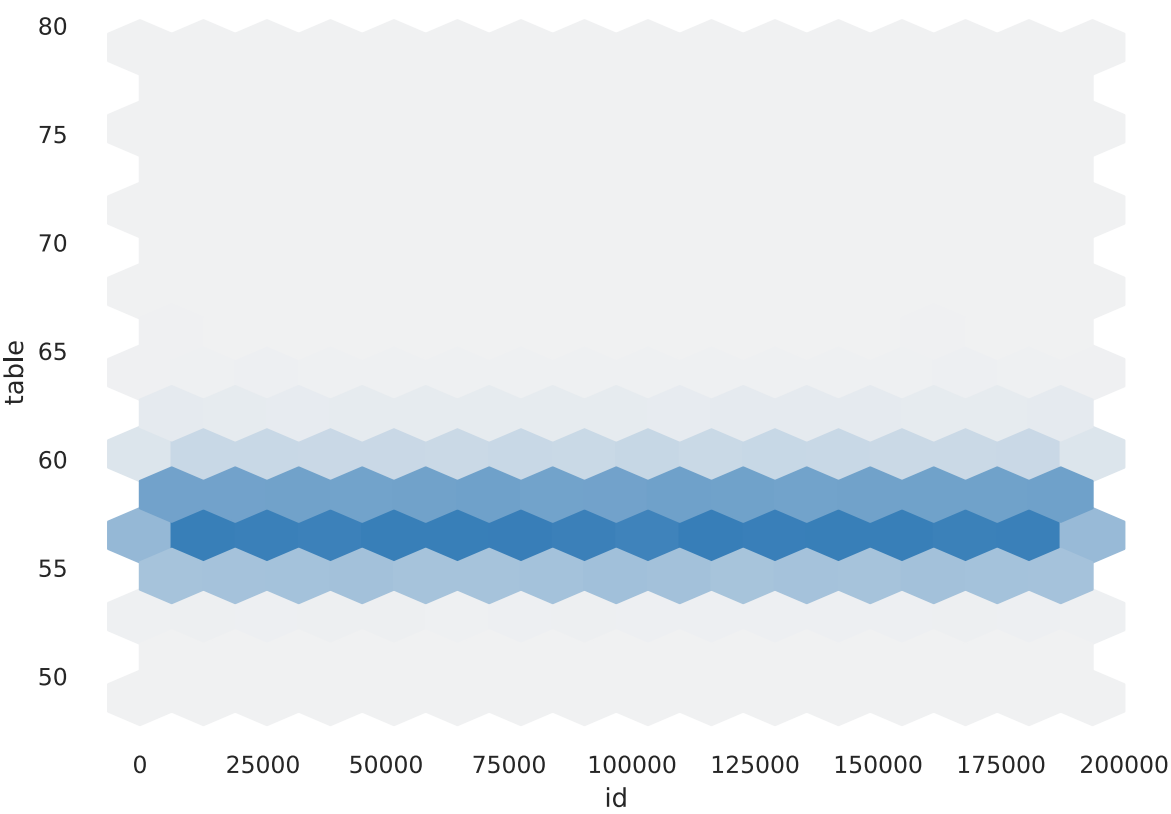
Interactions

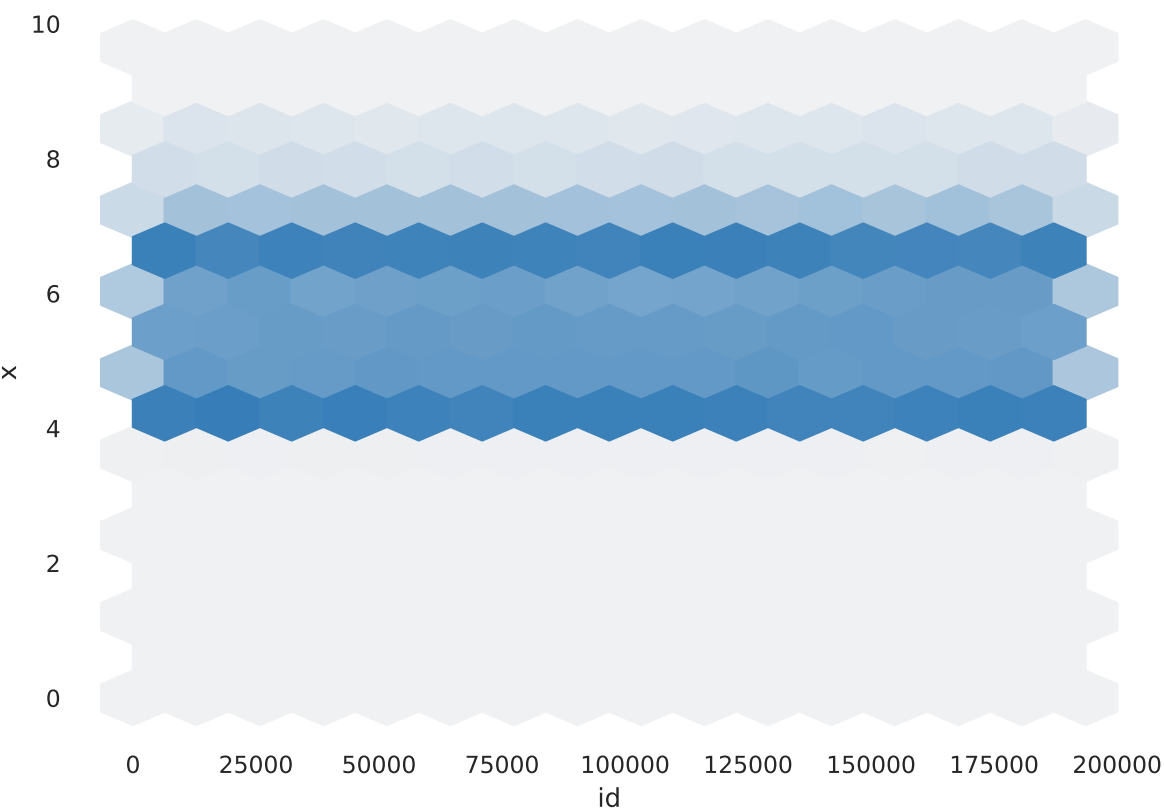


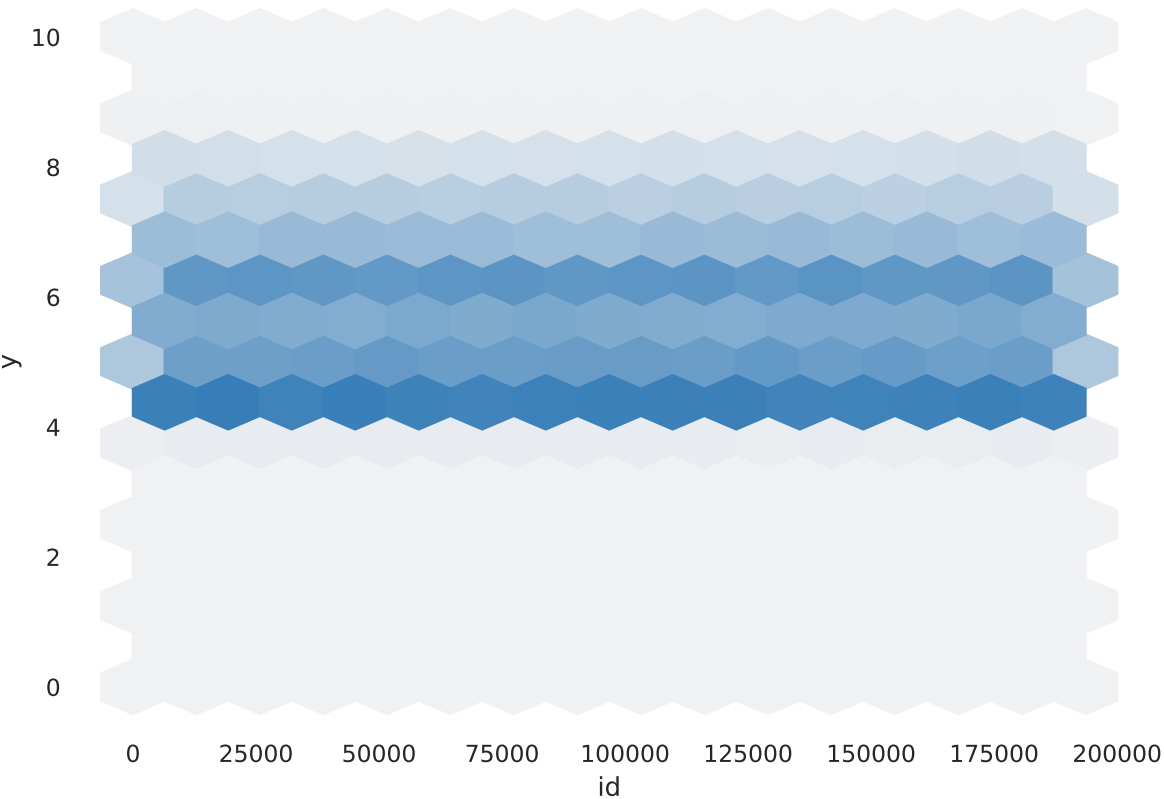


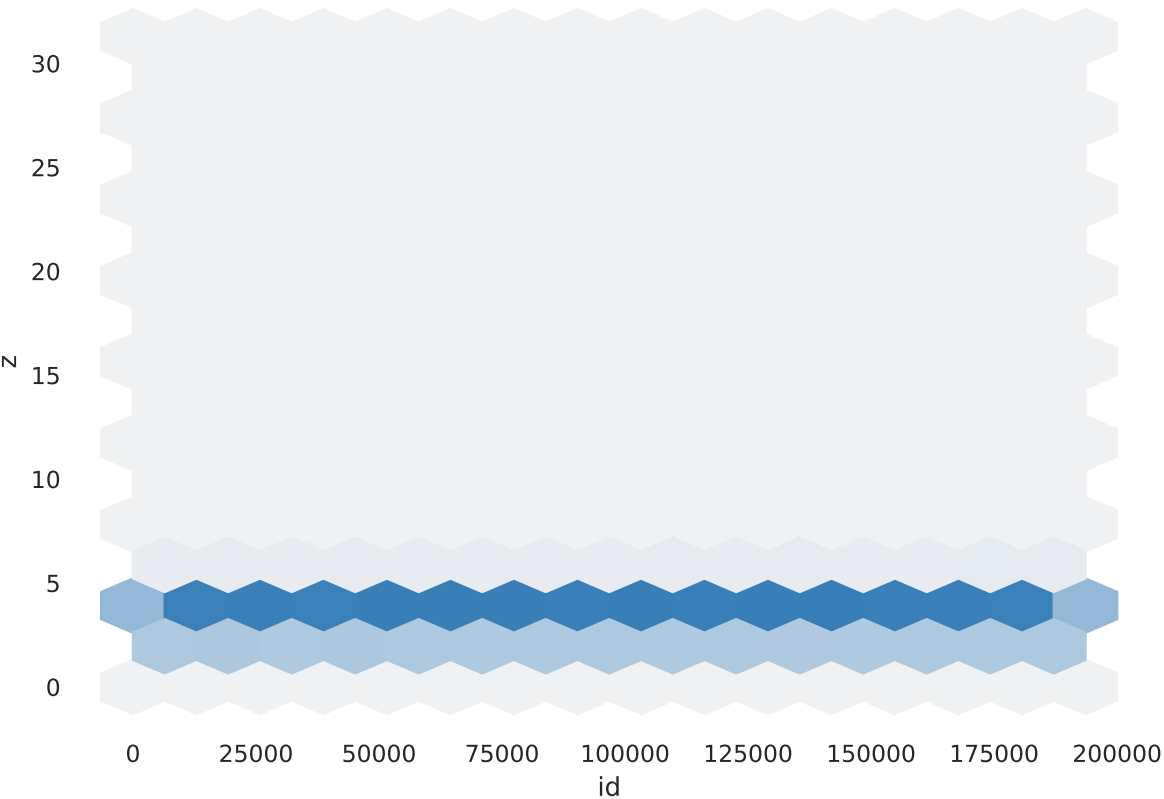


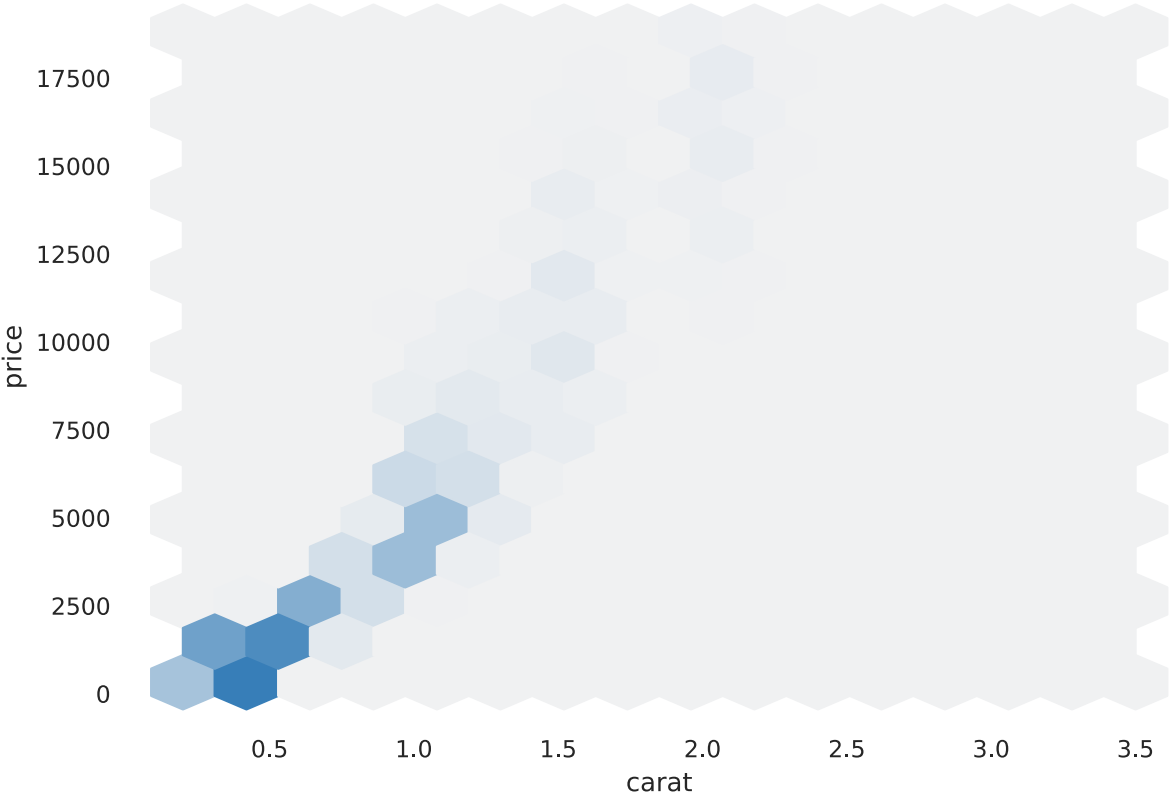


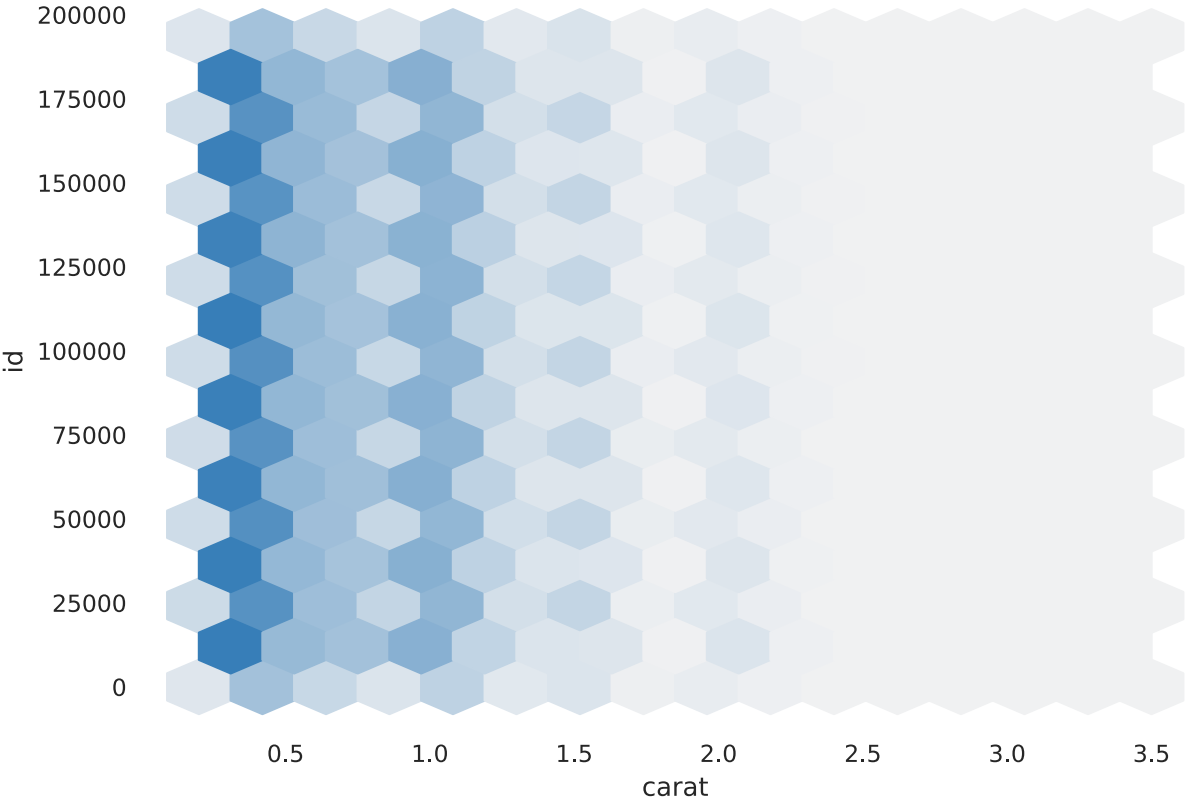


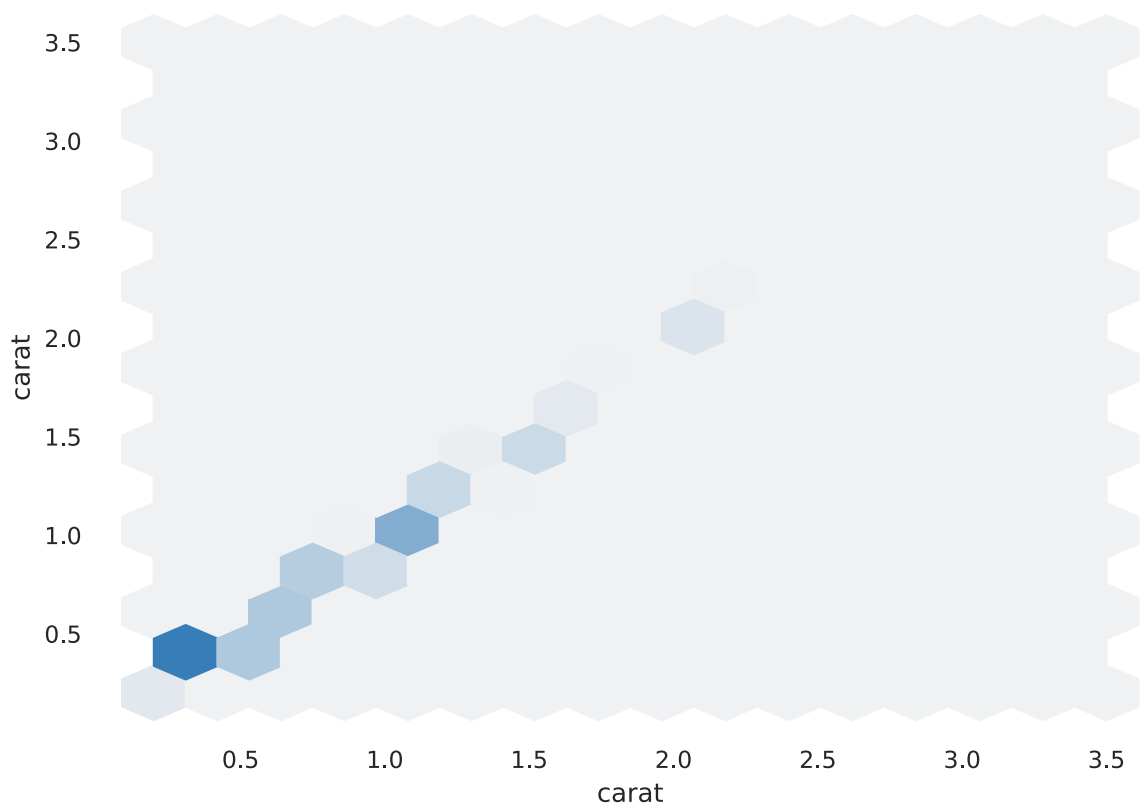


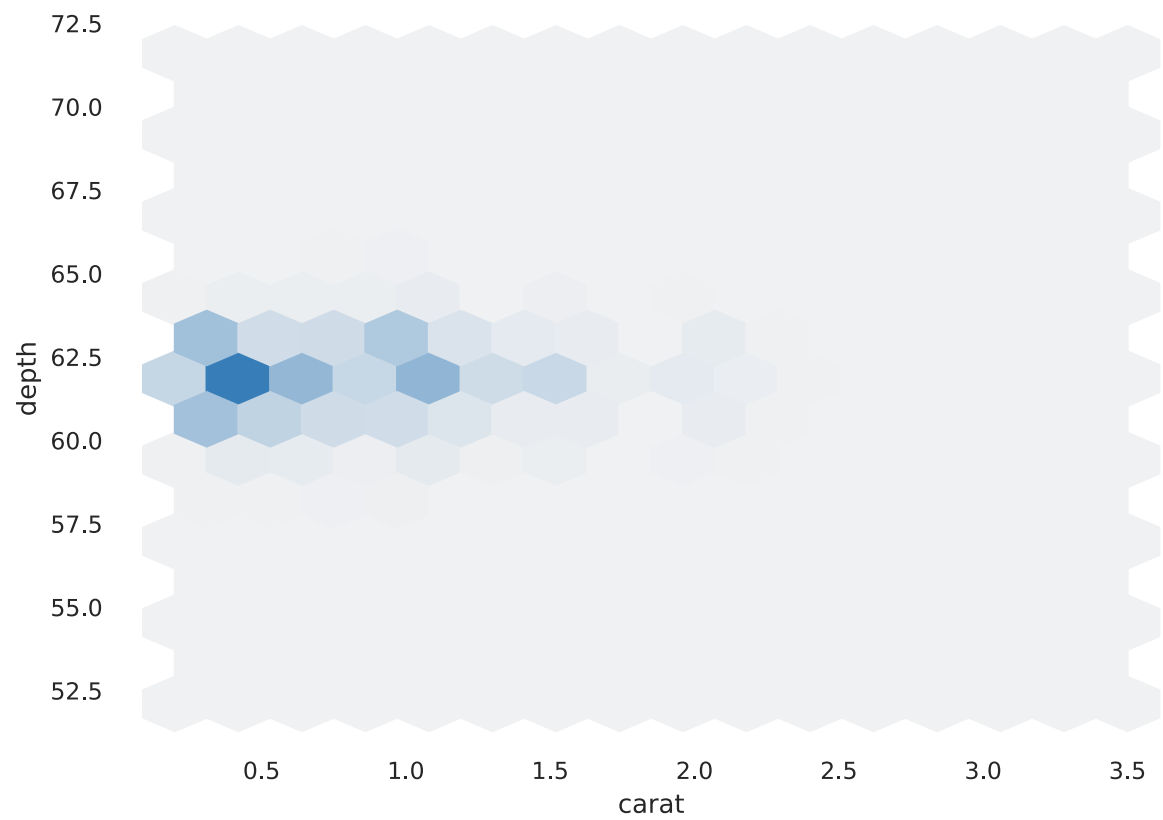


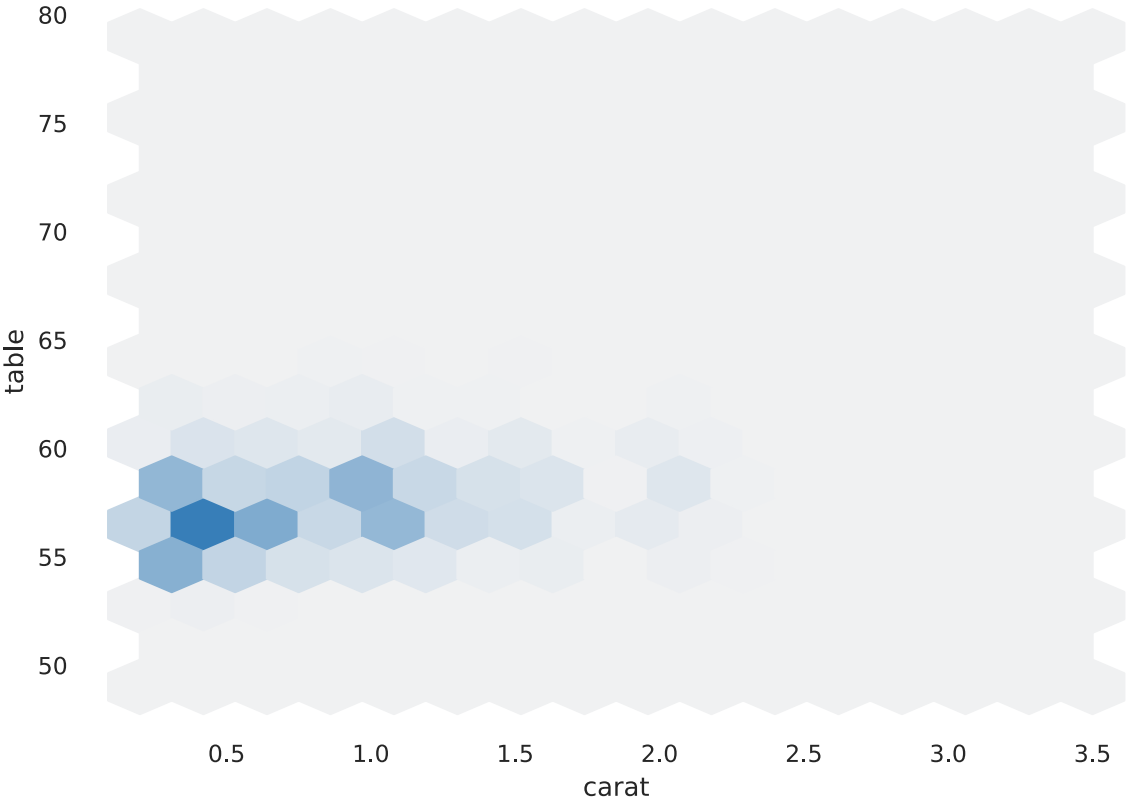


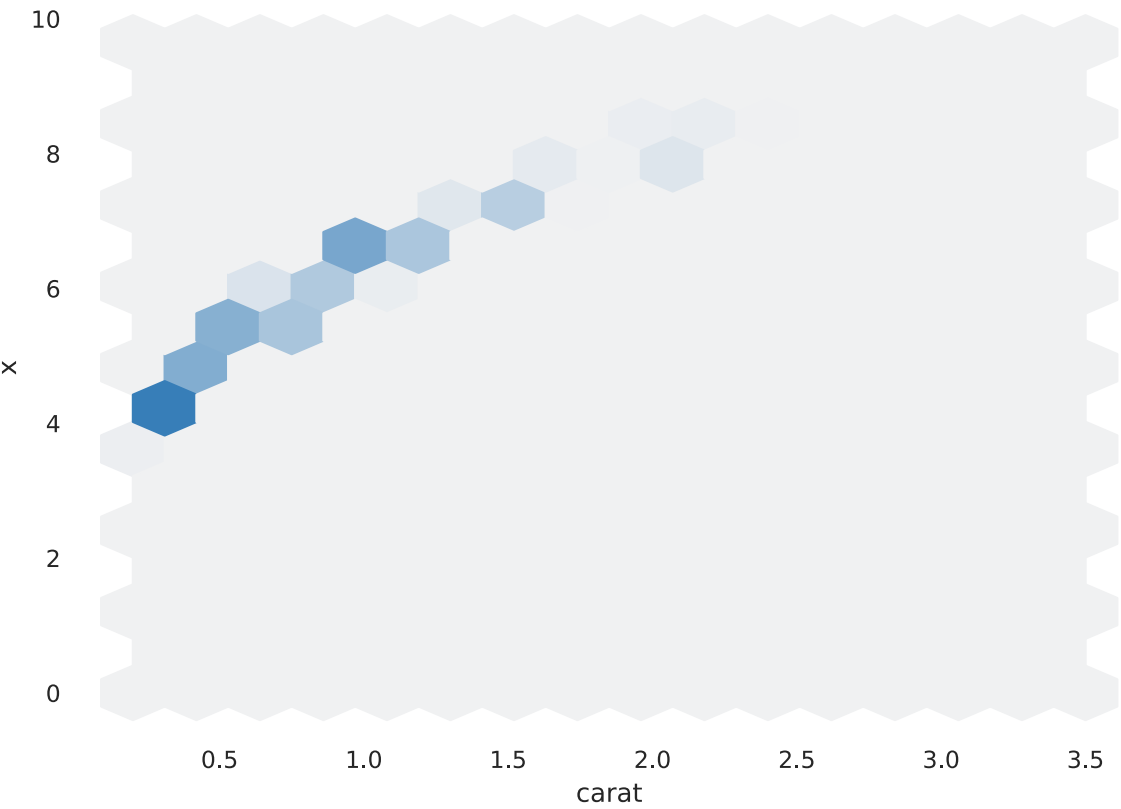


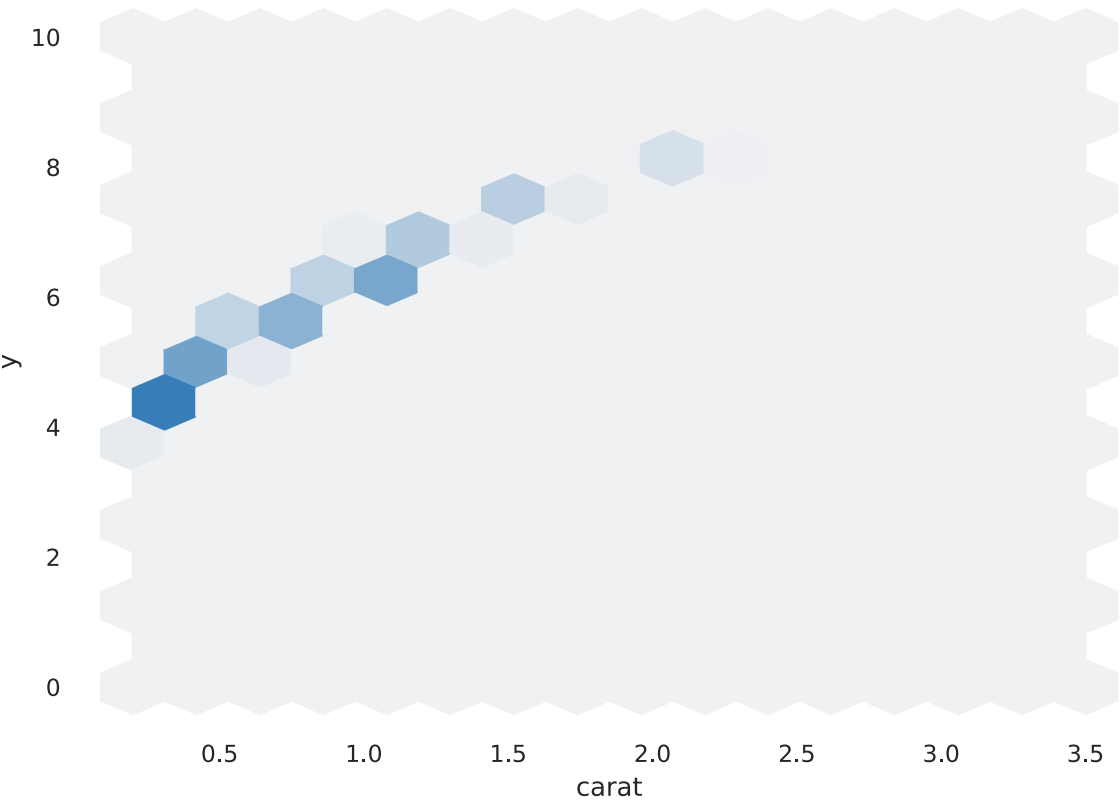


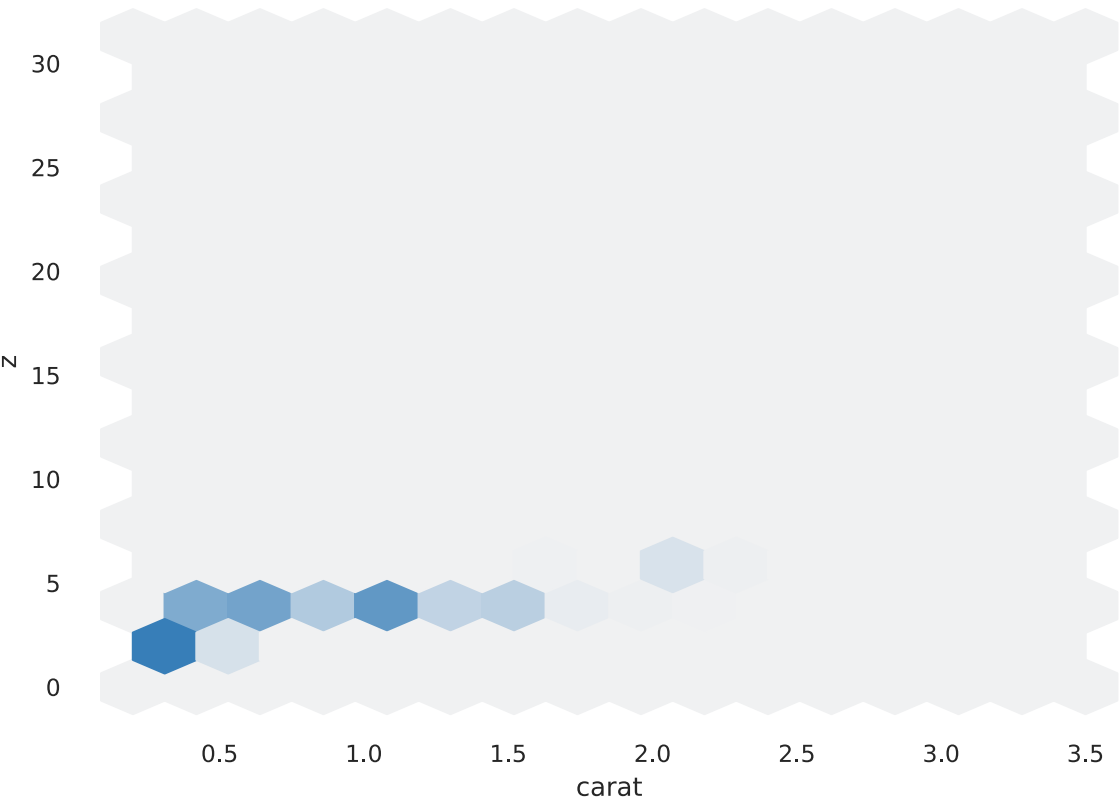


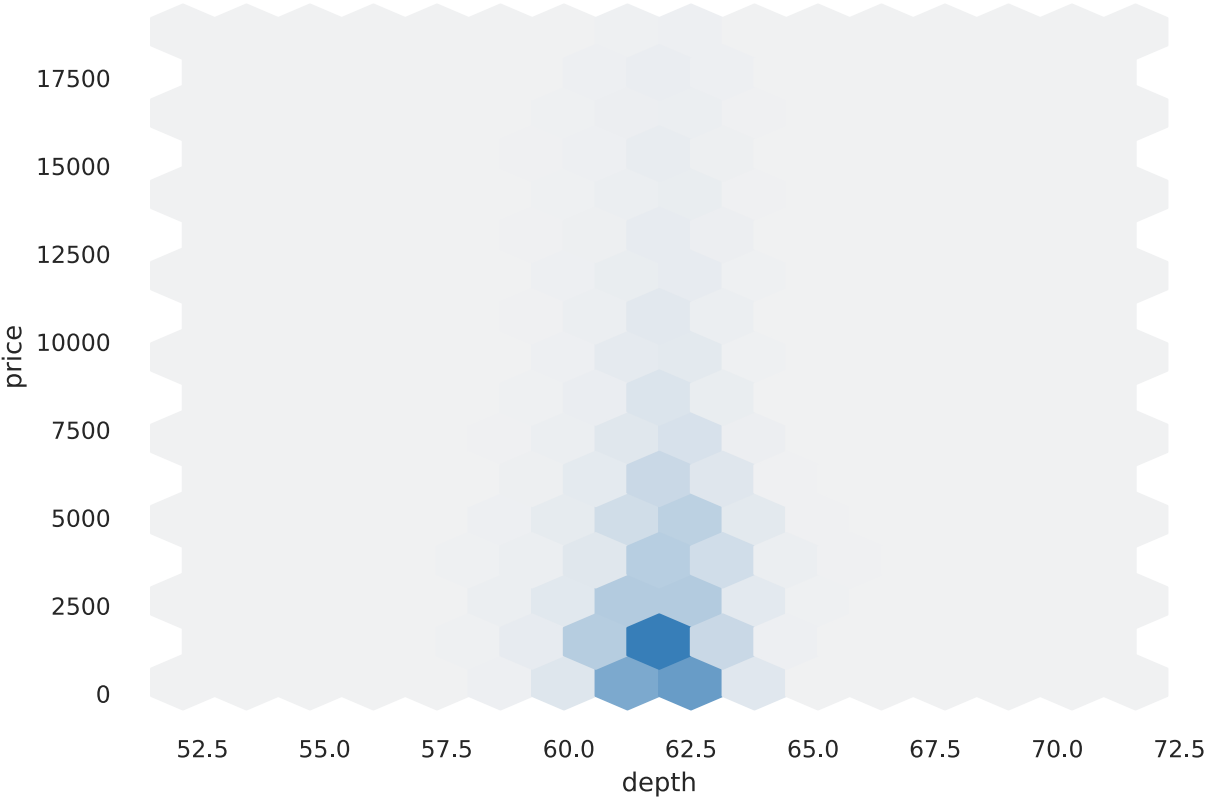


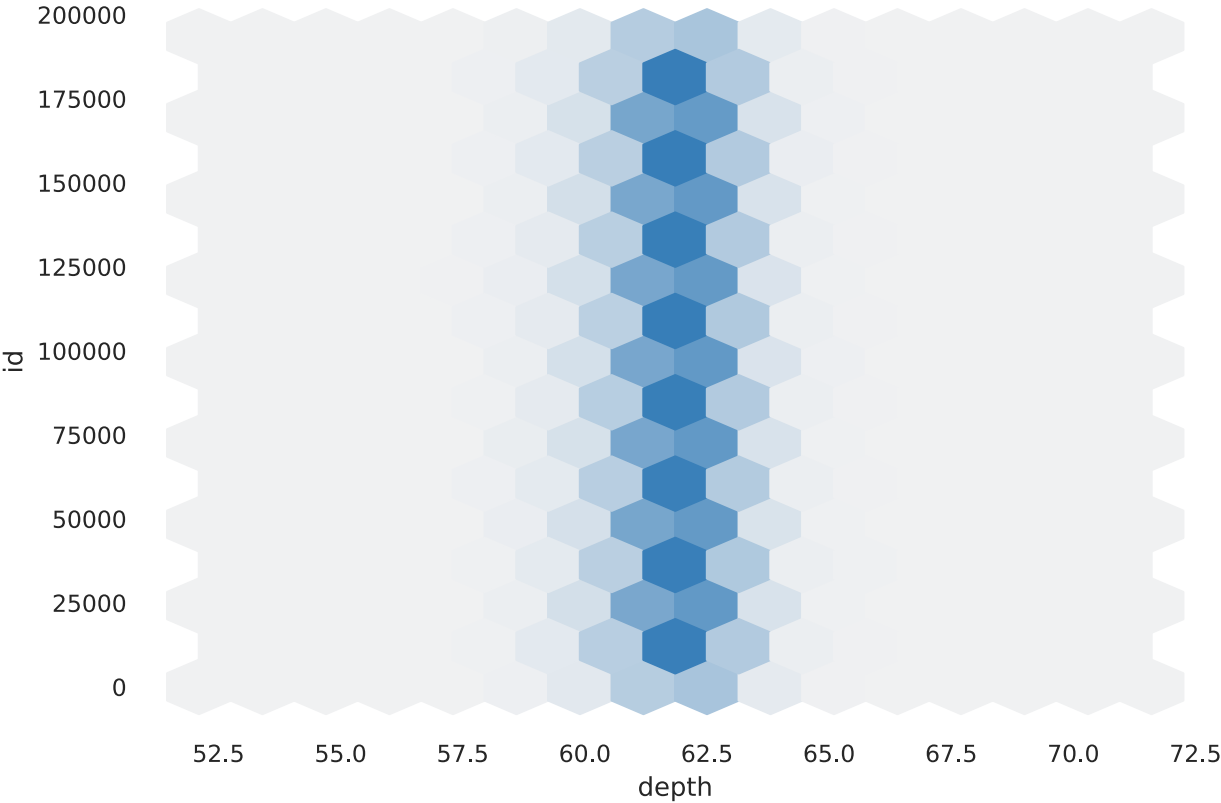


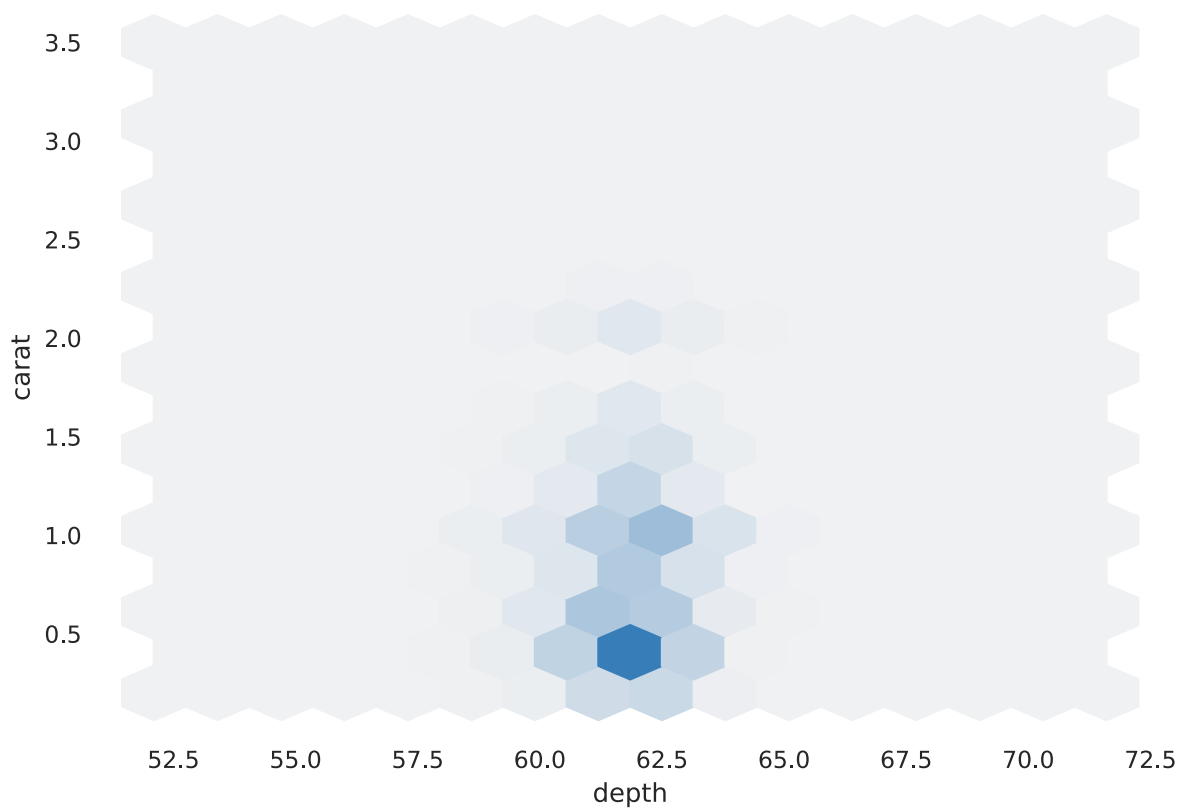


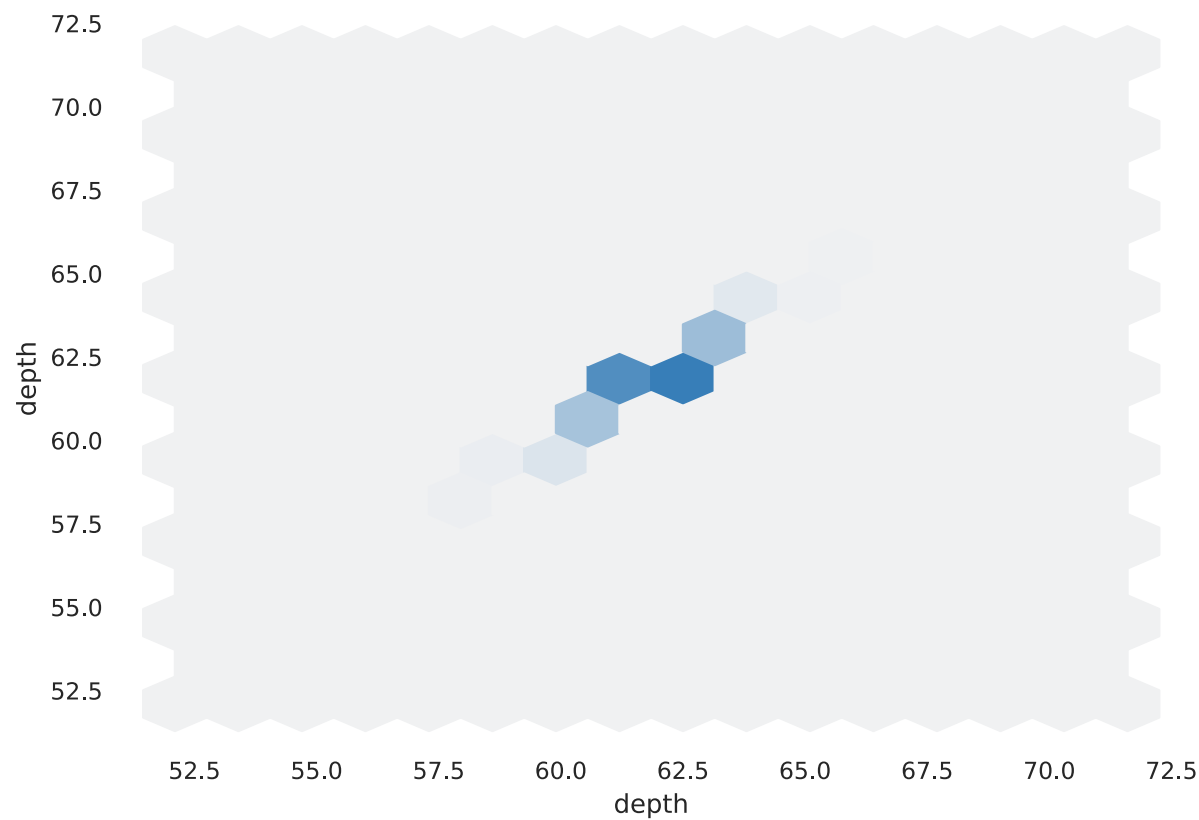


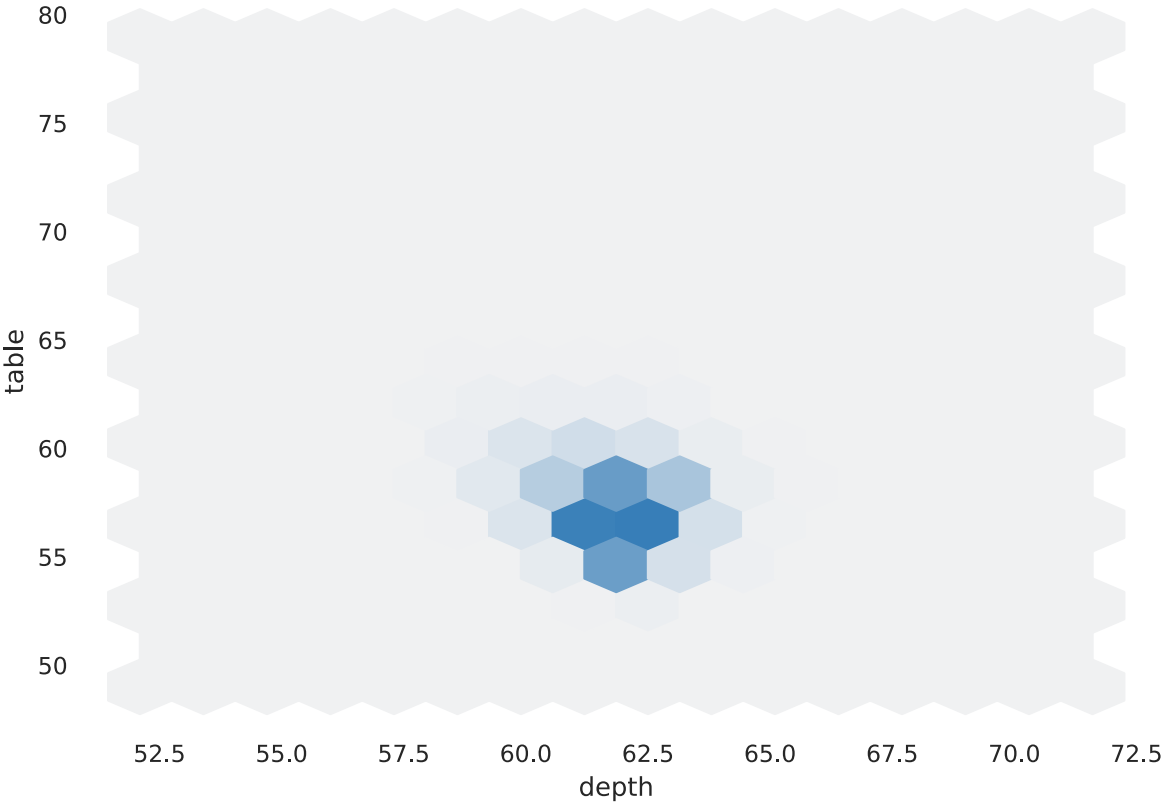


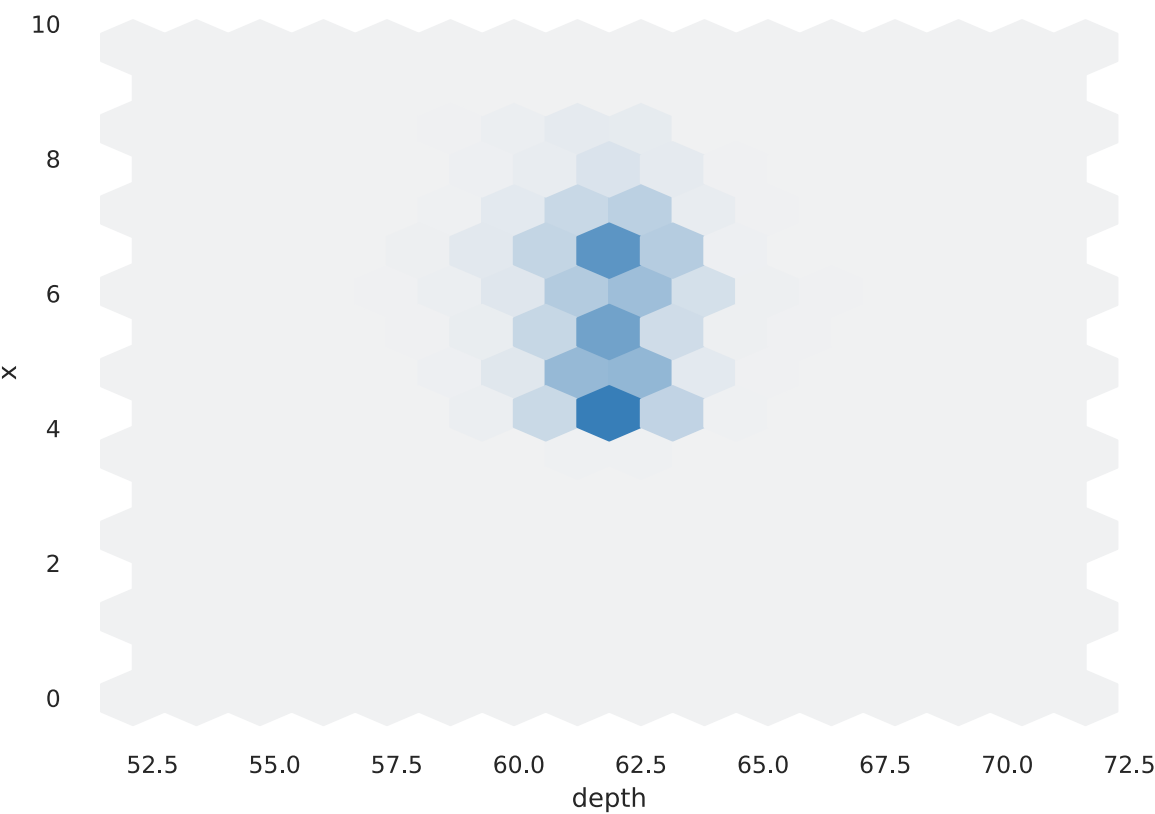


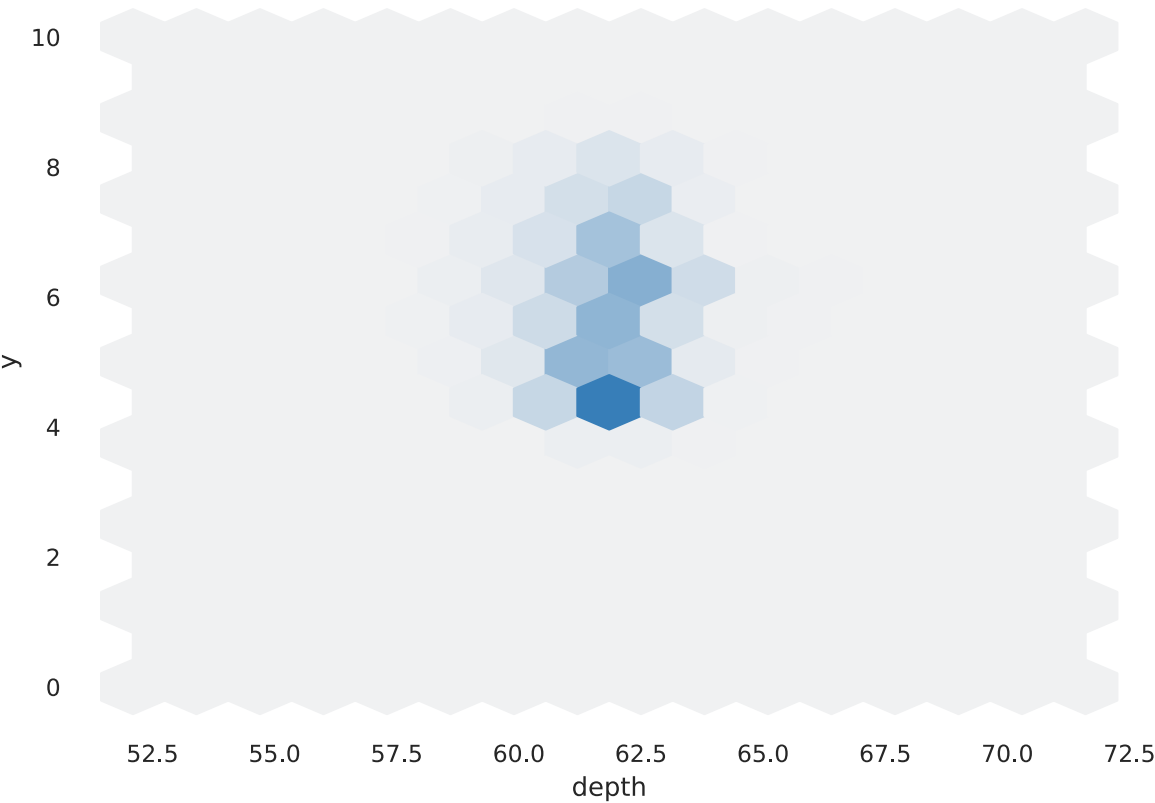


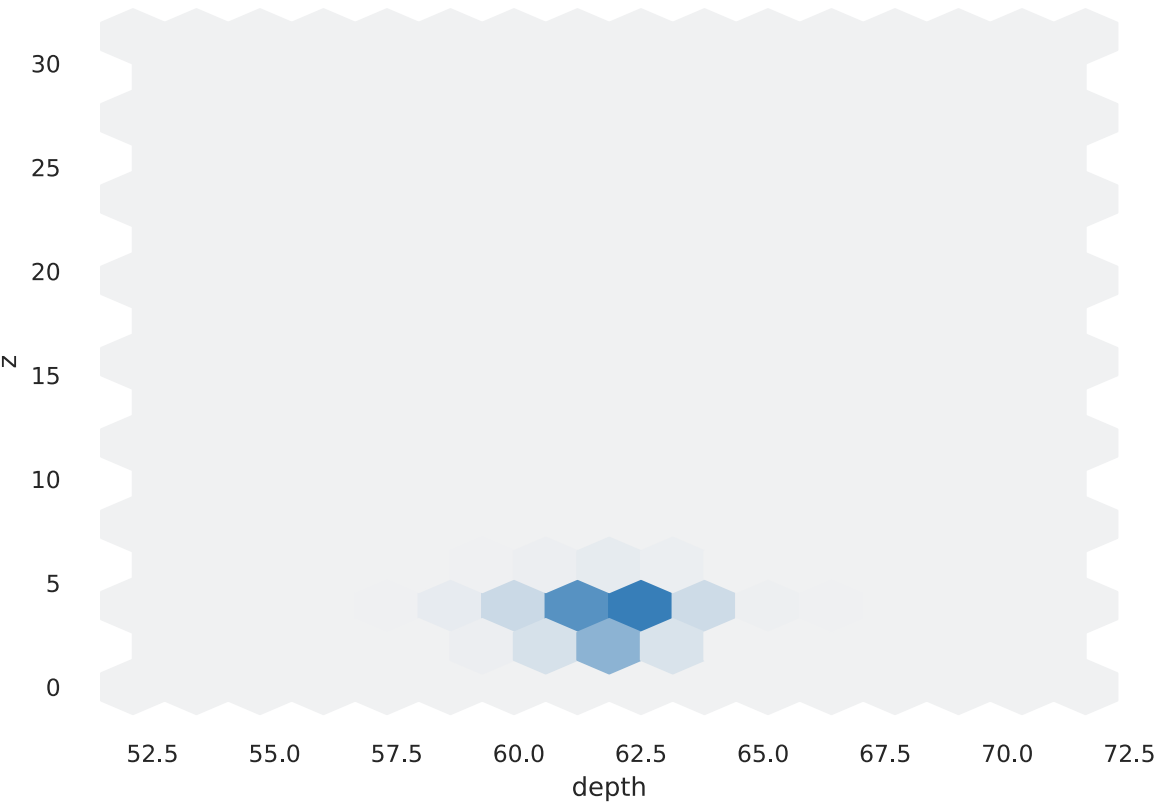


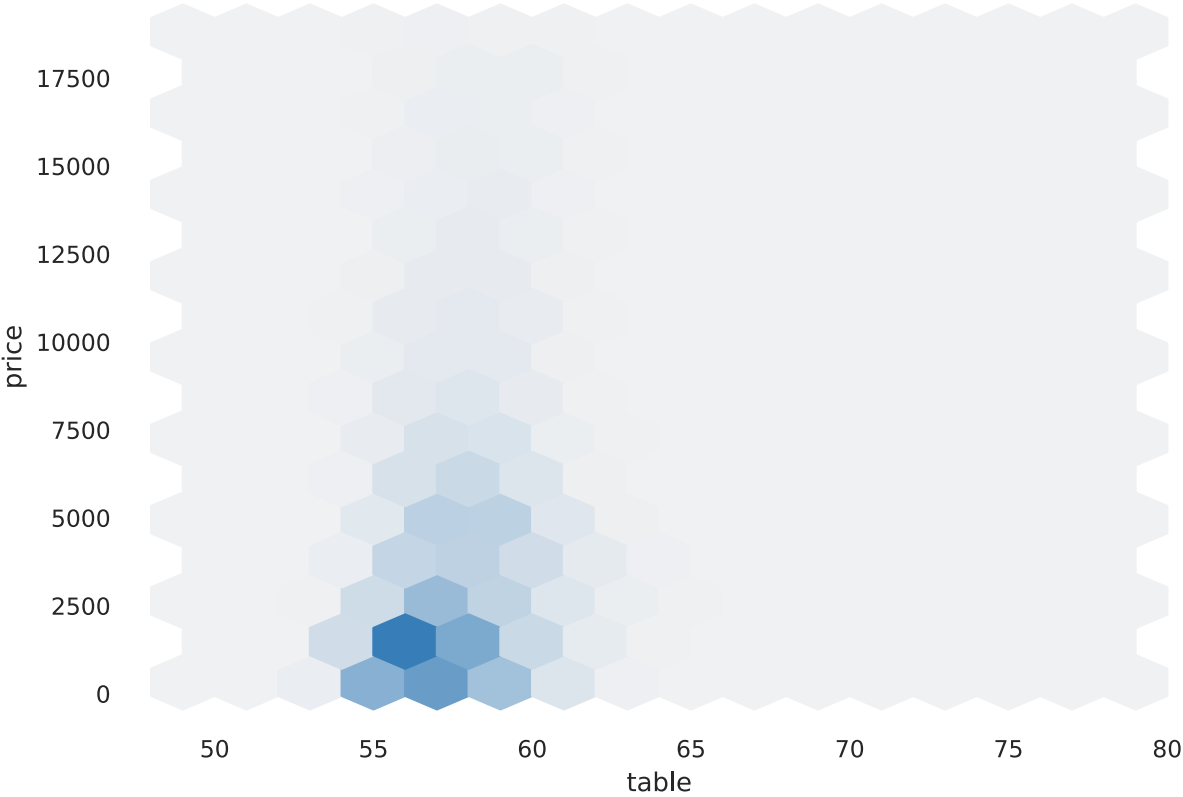


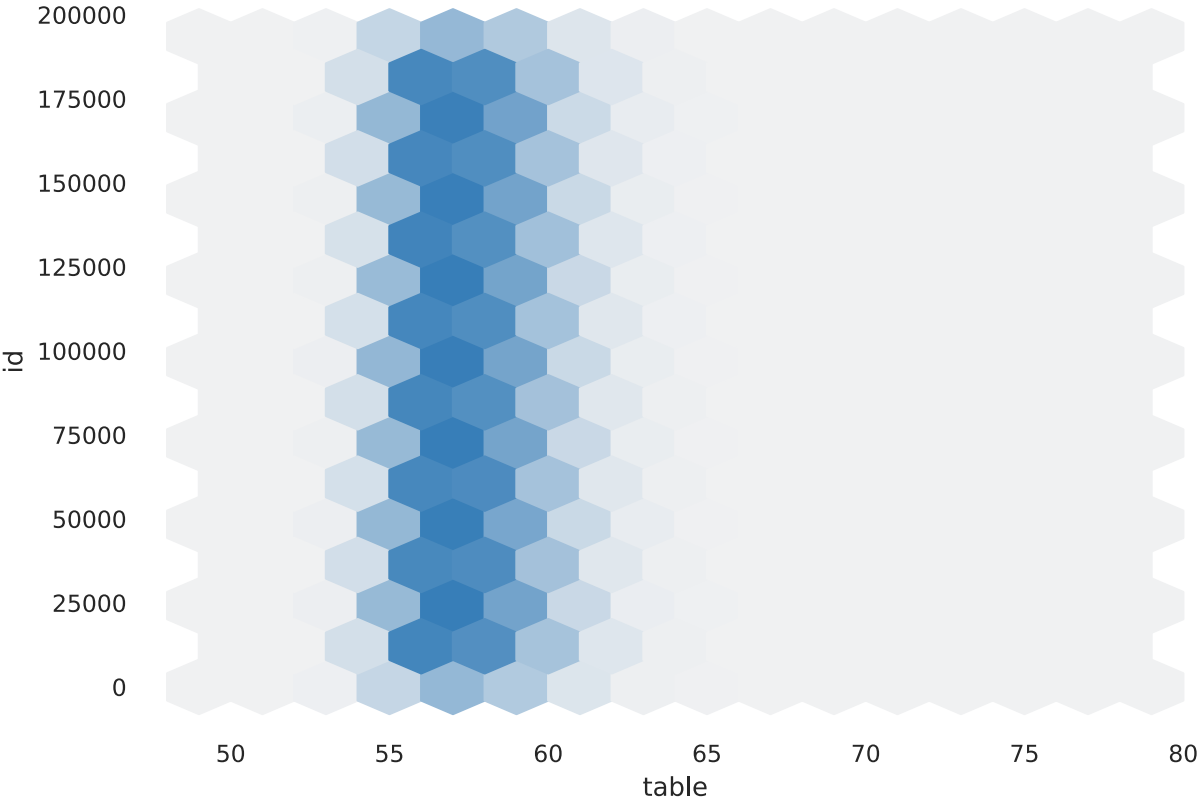


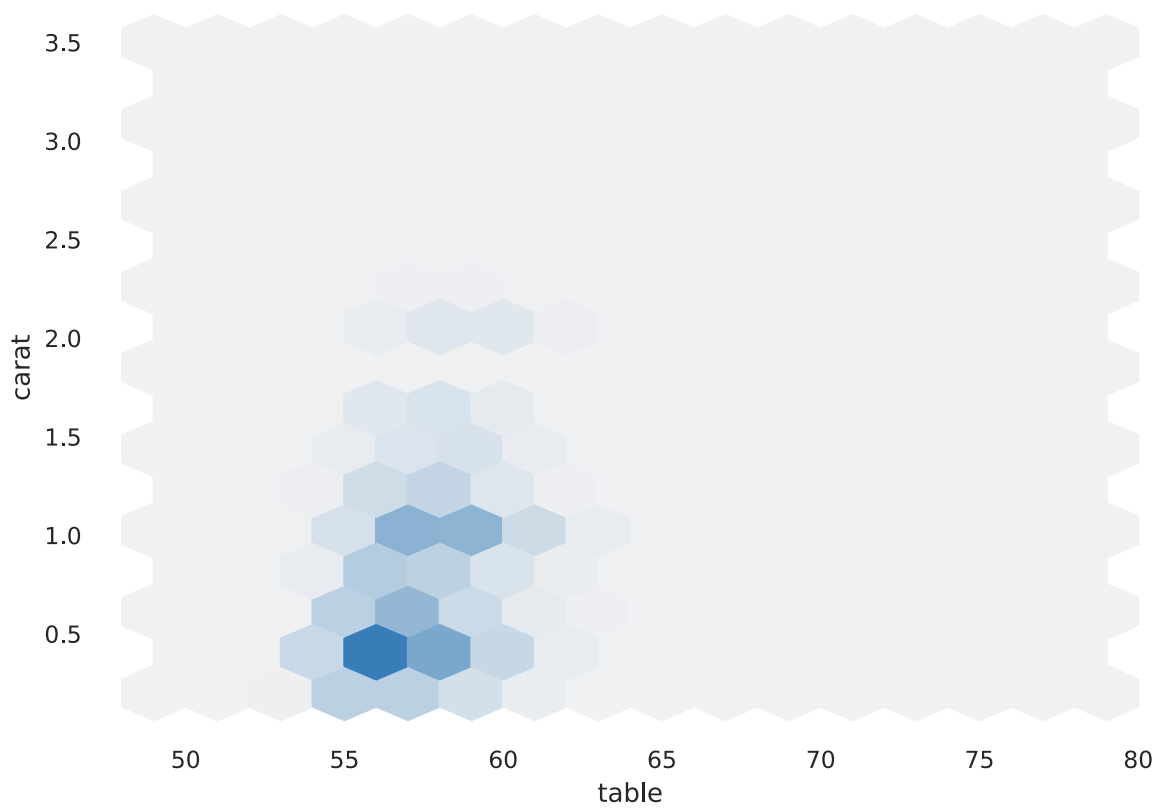


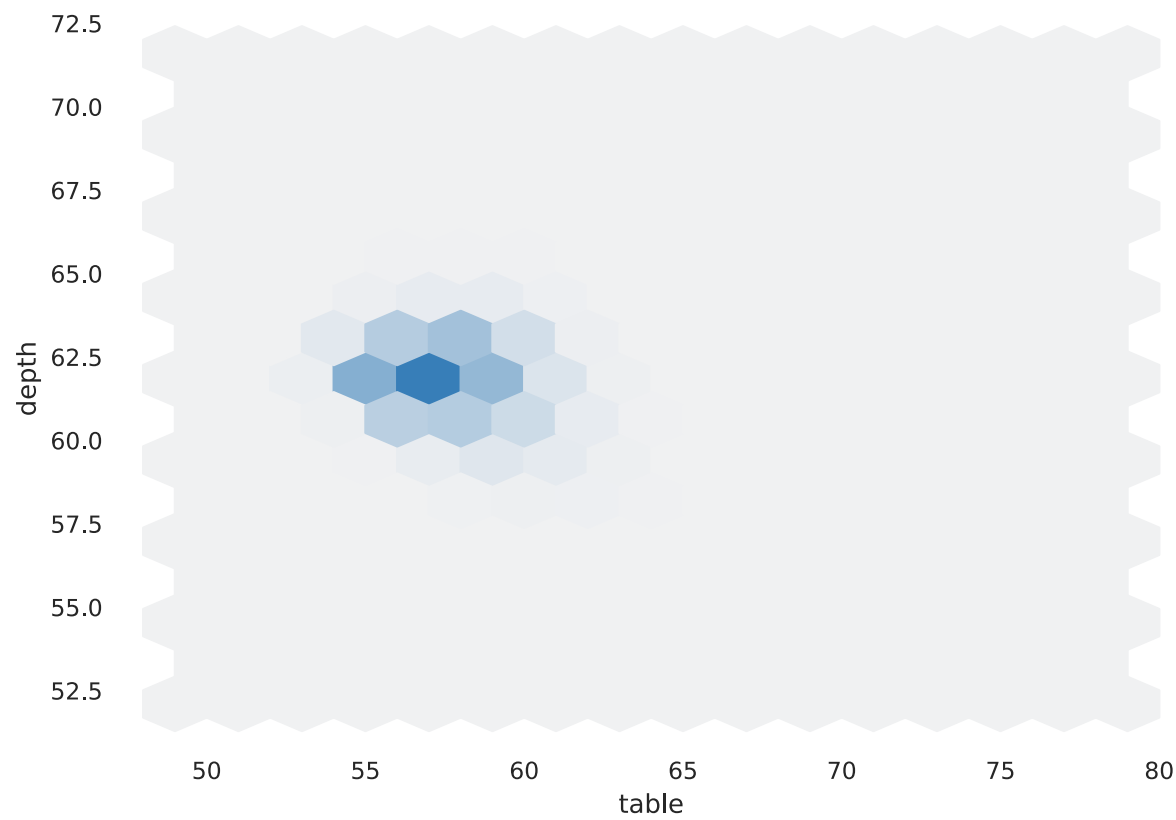


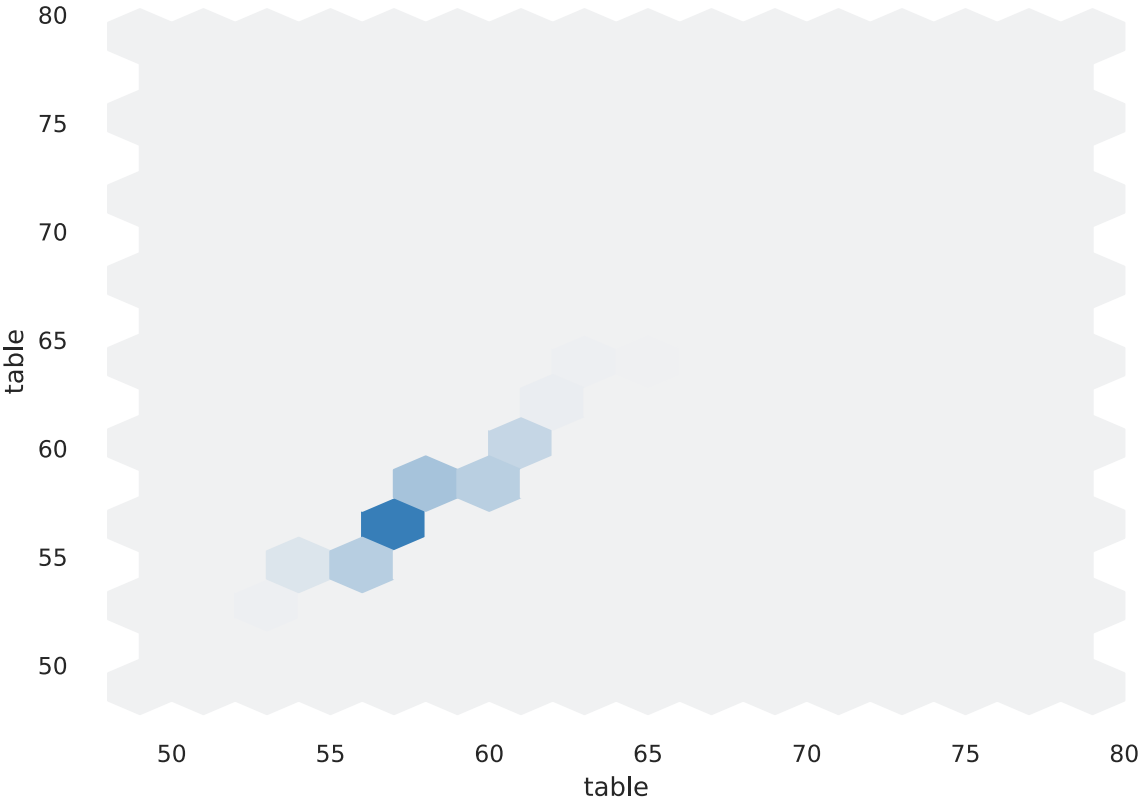


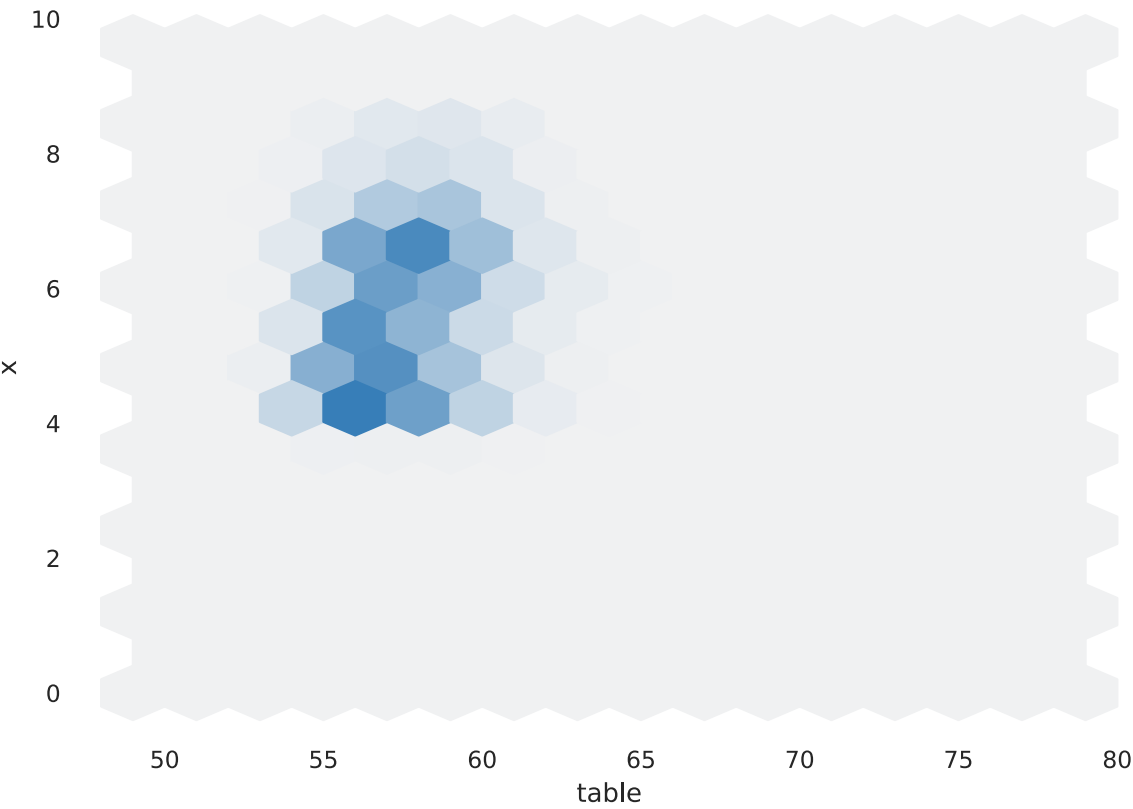


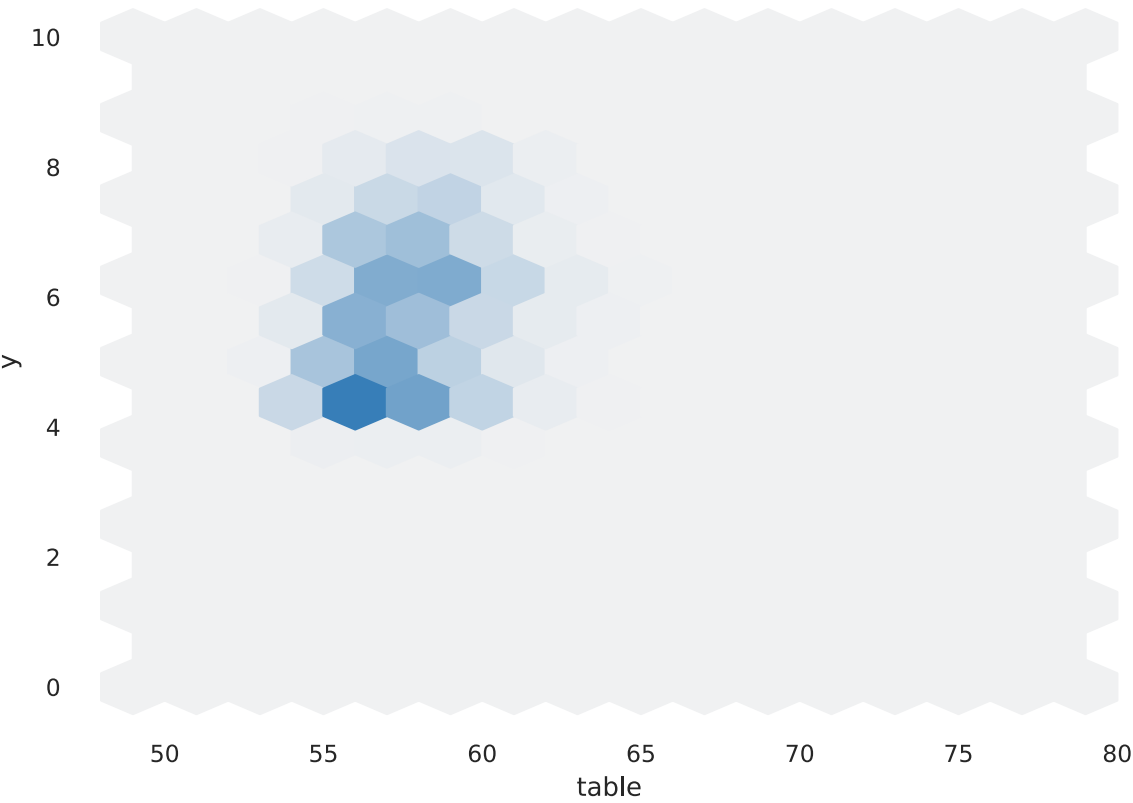


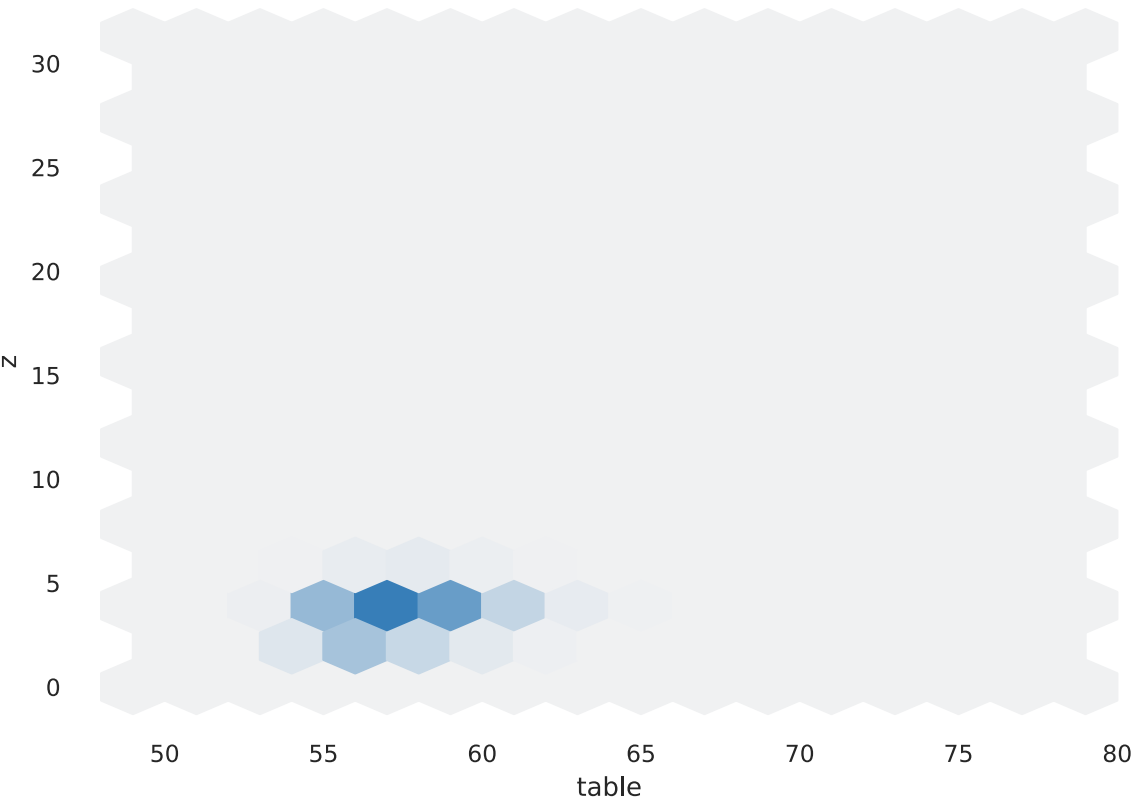


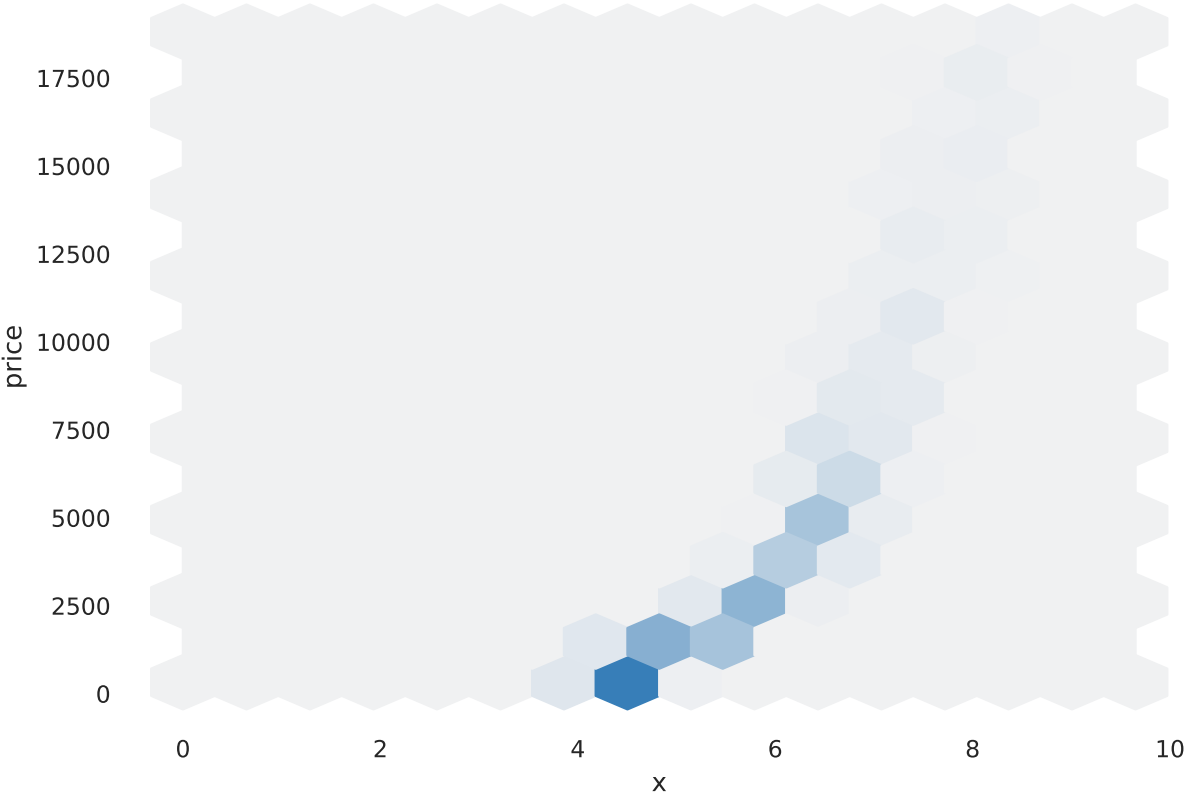


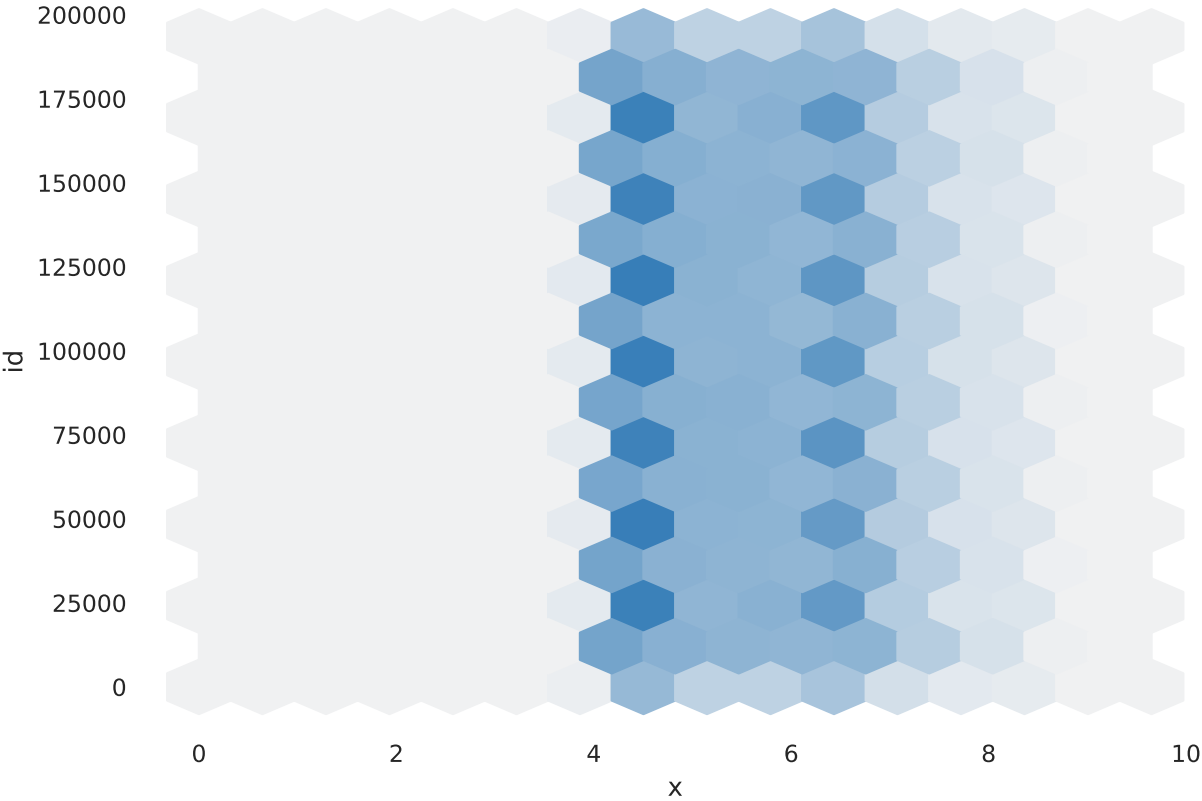


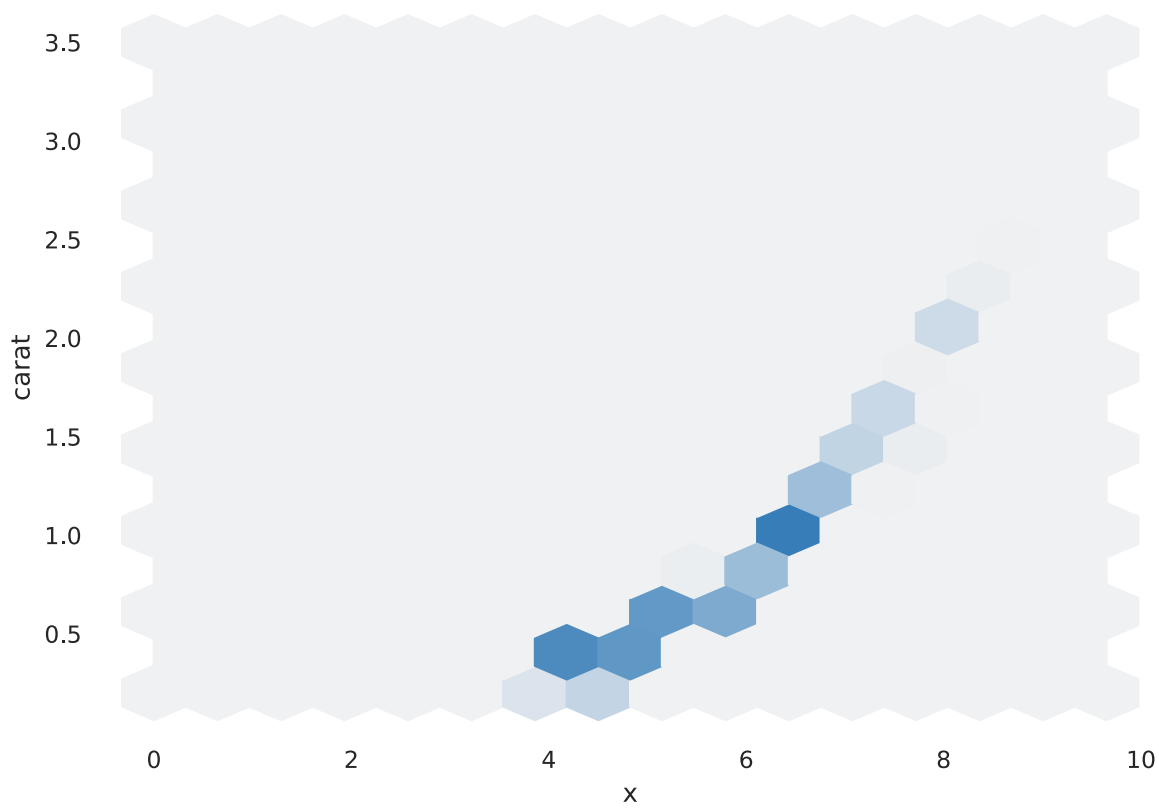


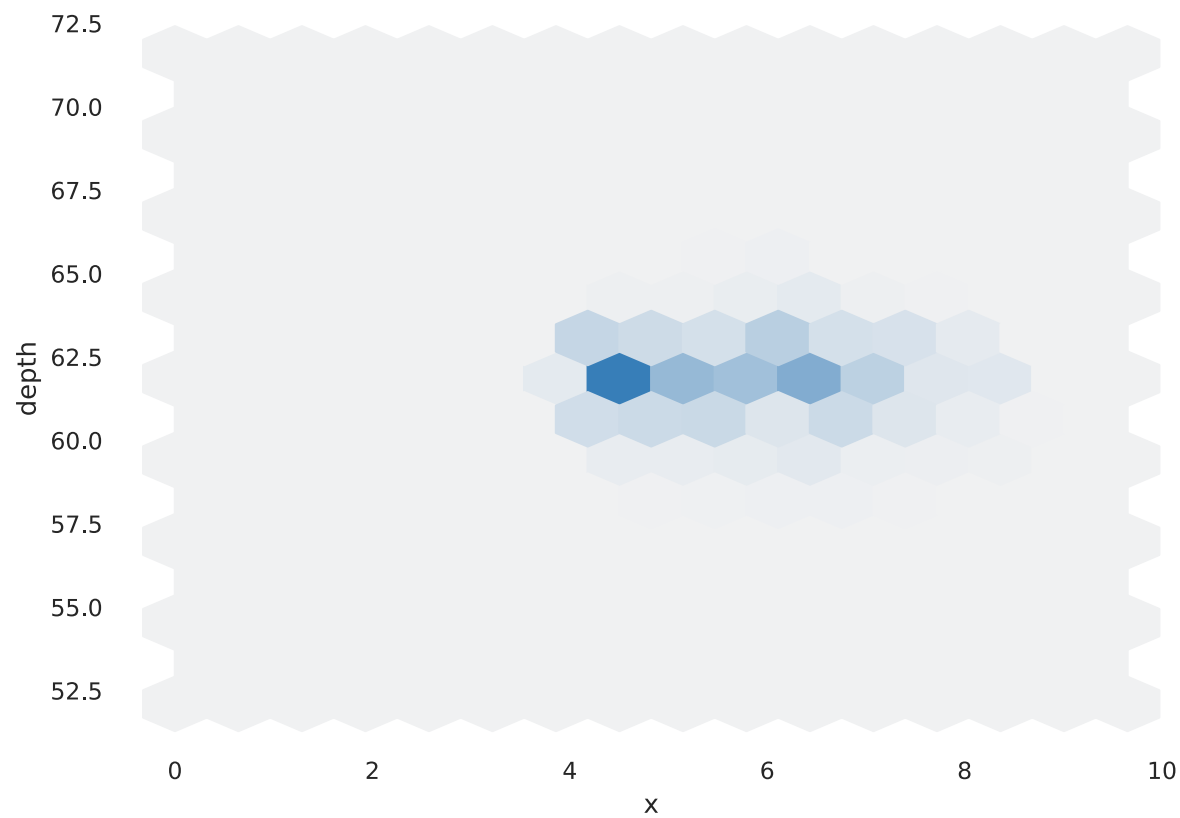


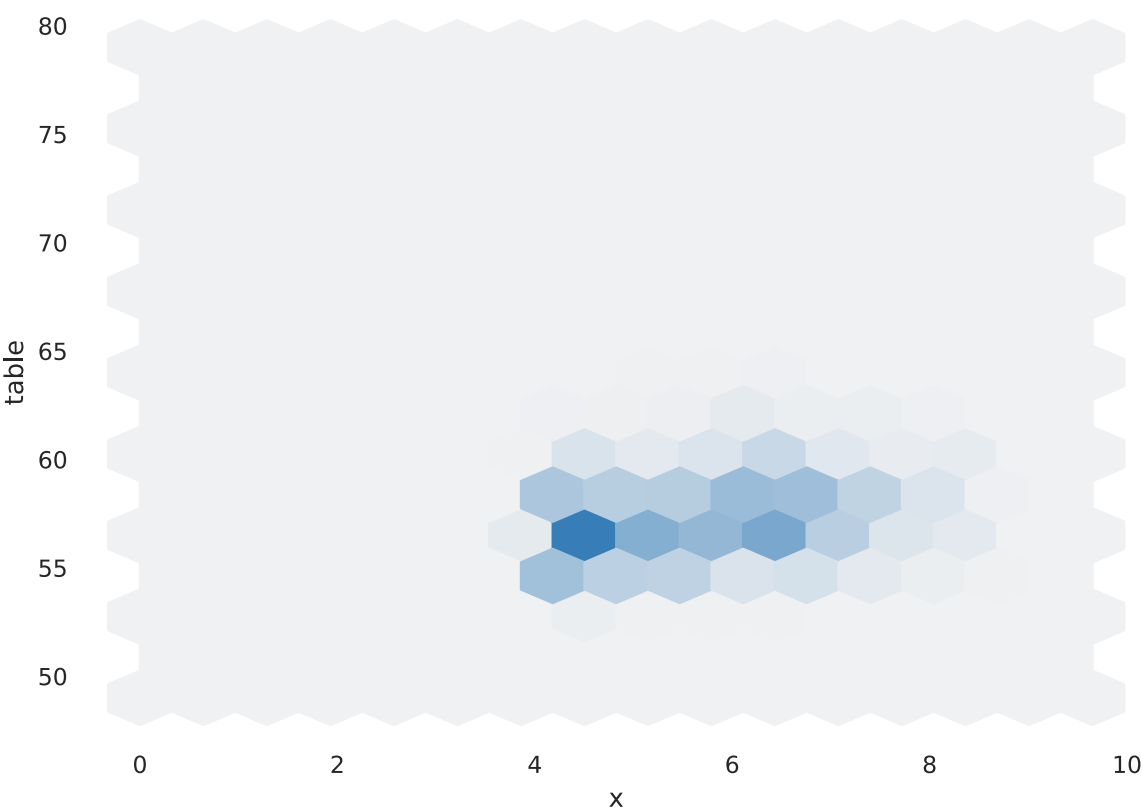


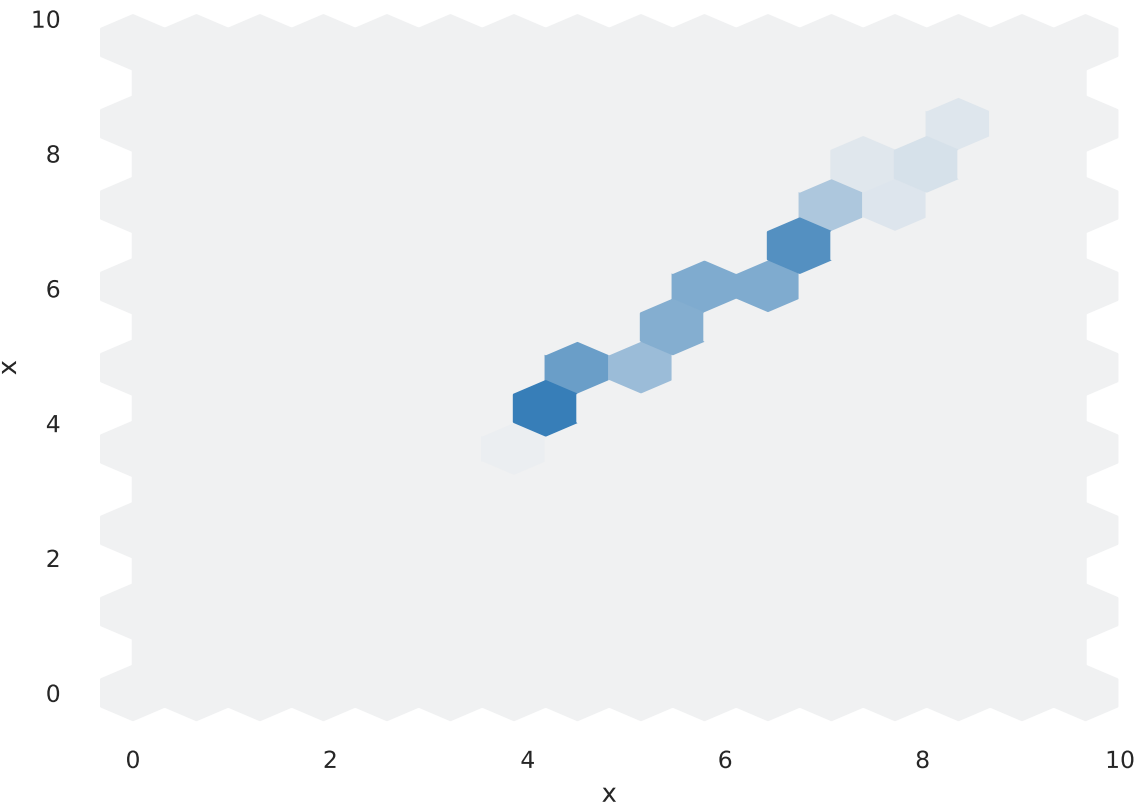


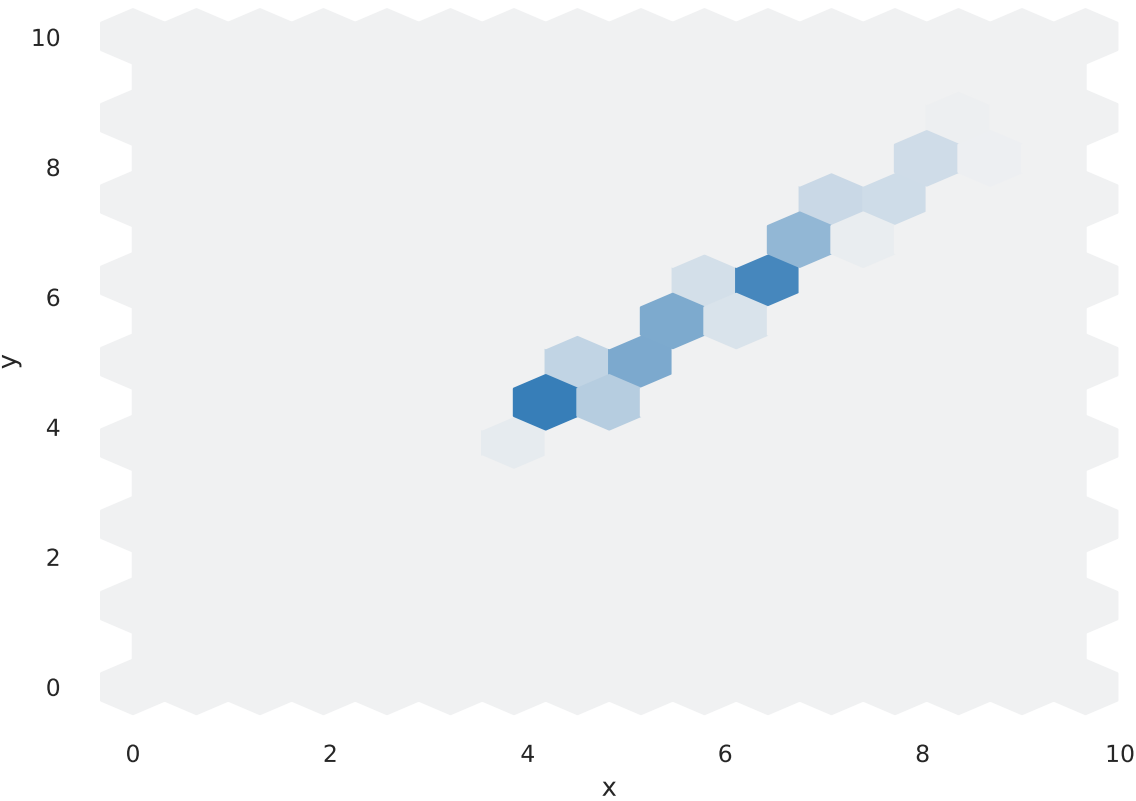


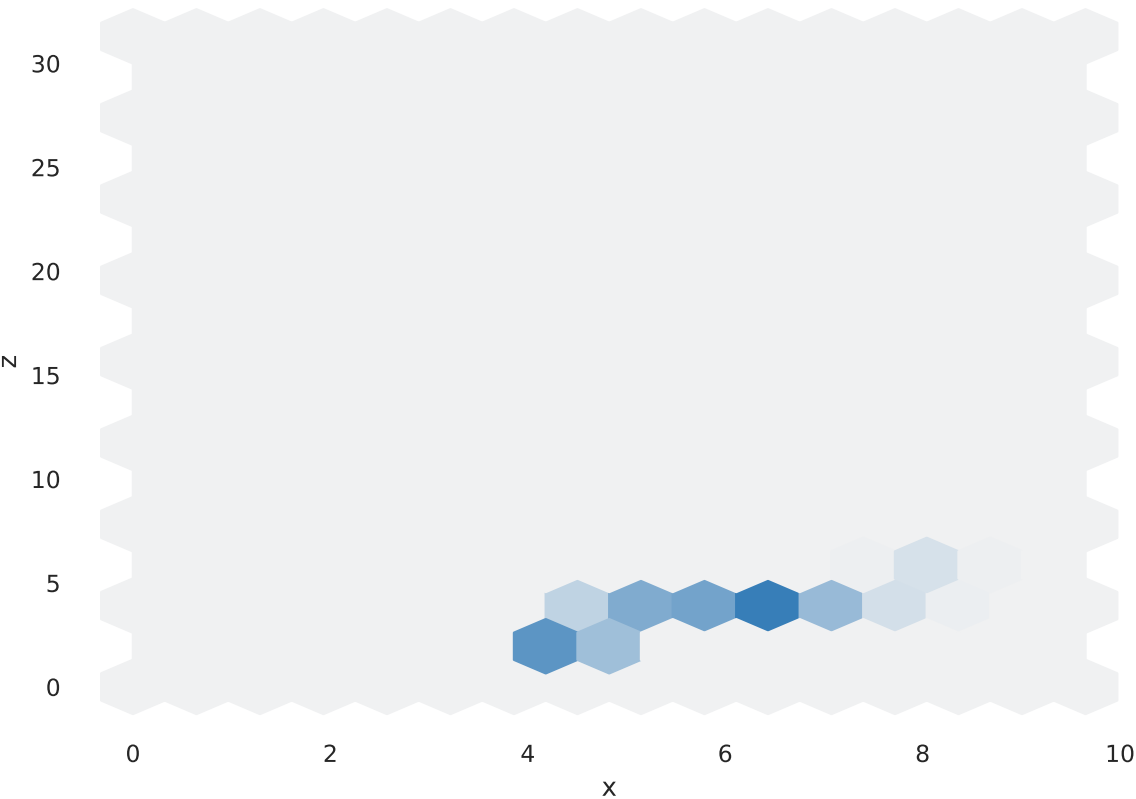


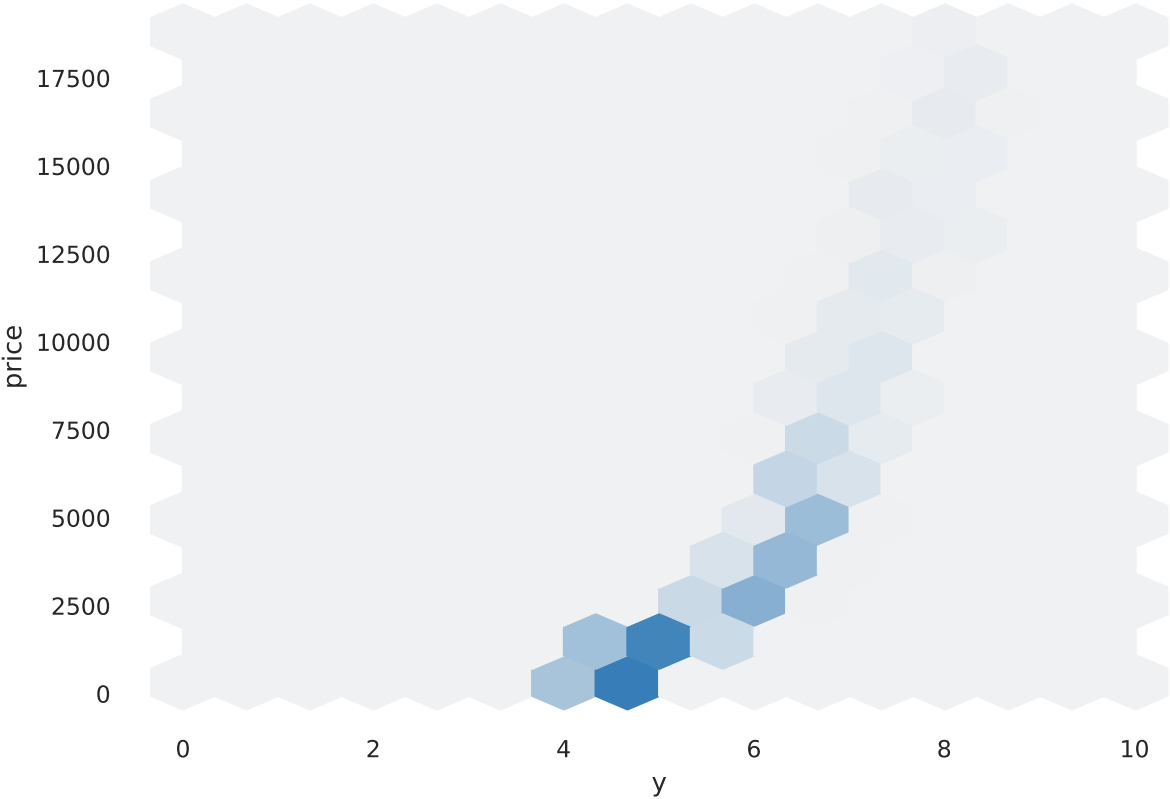


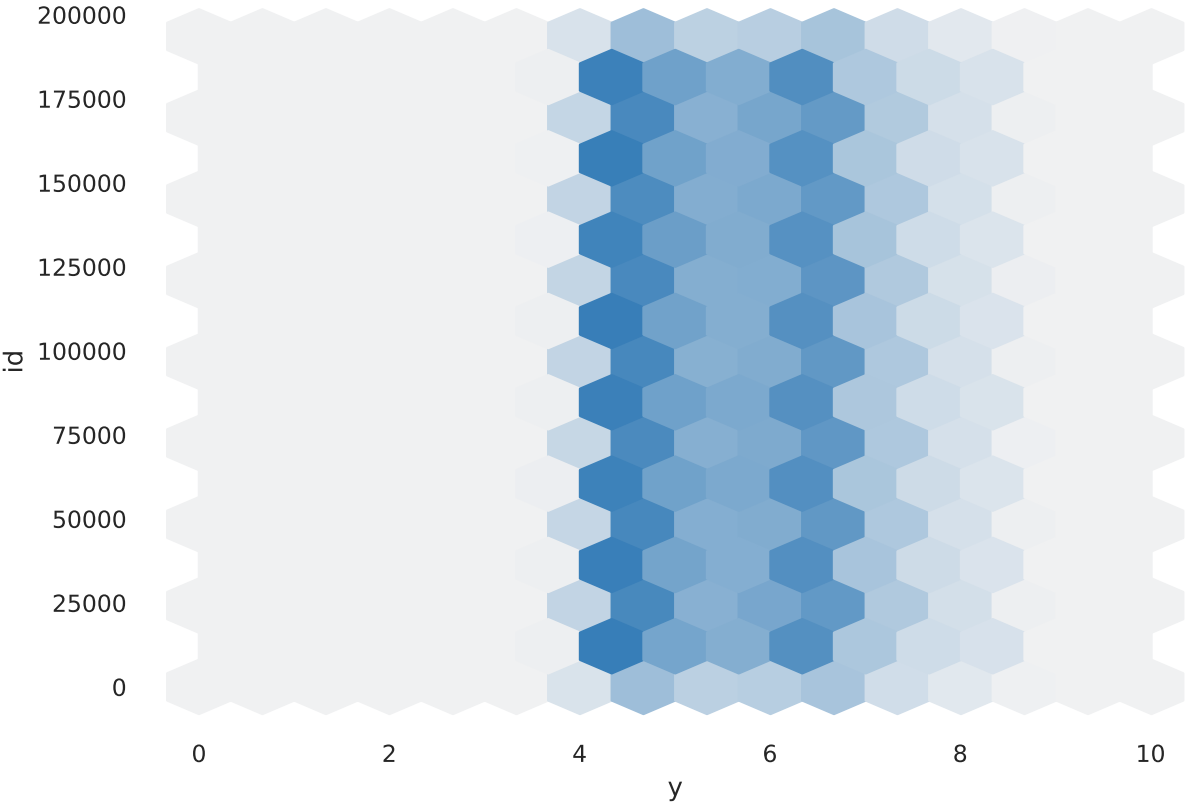


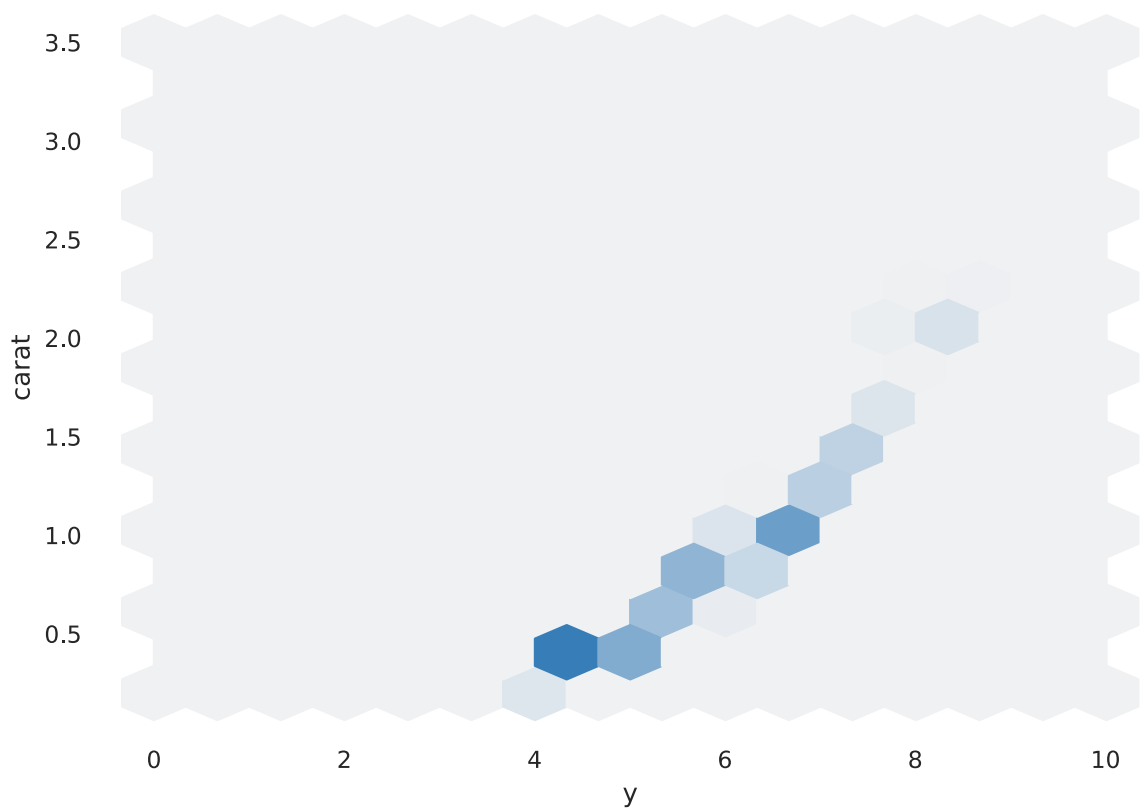


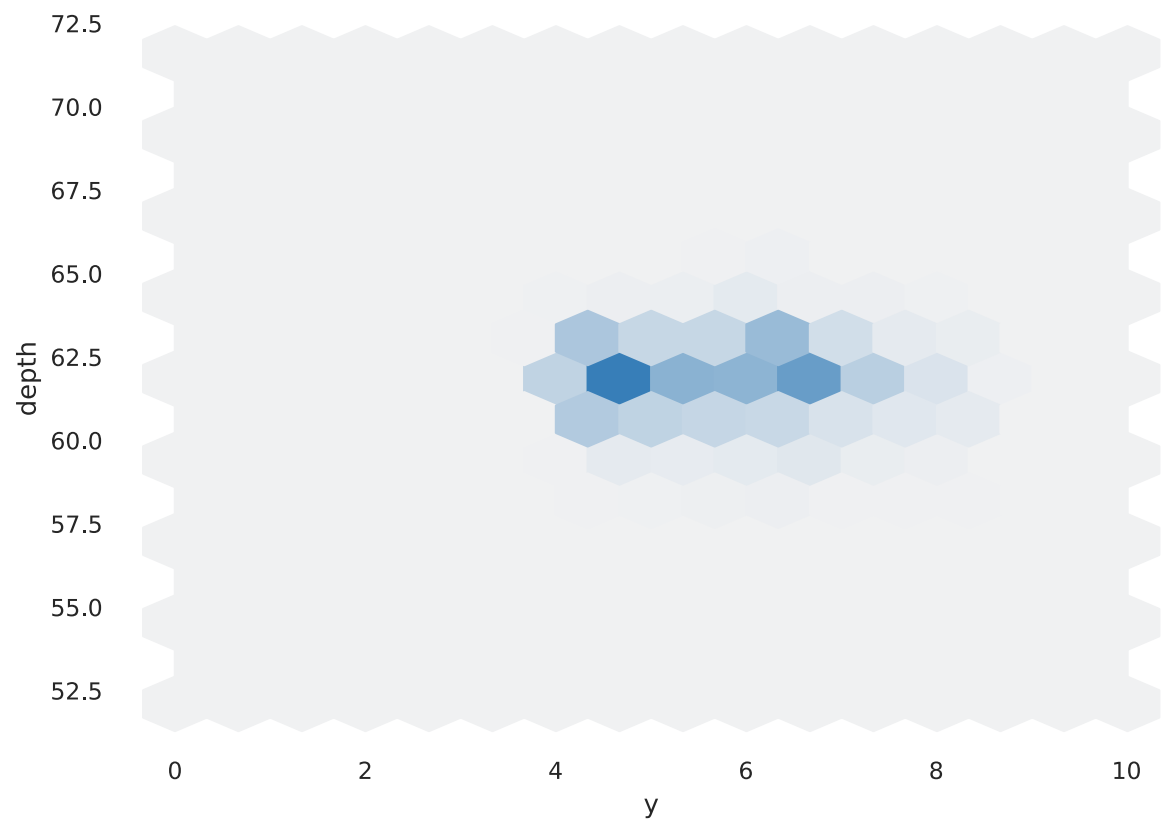


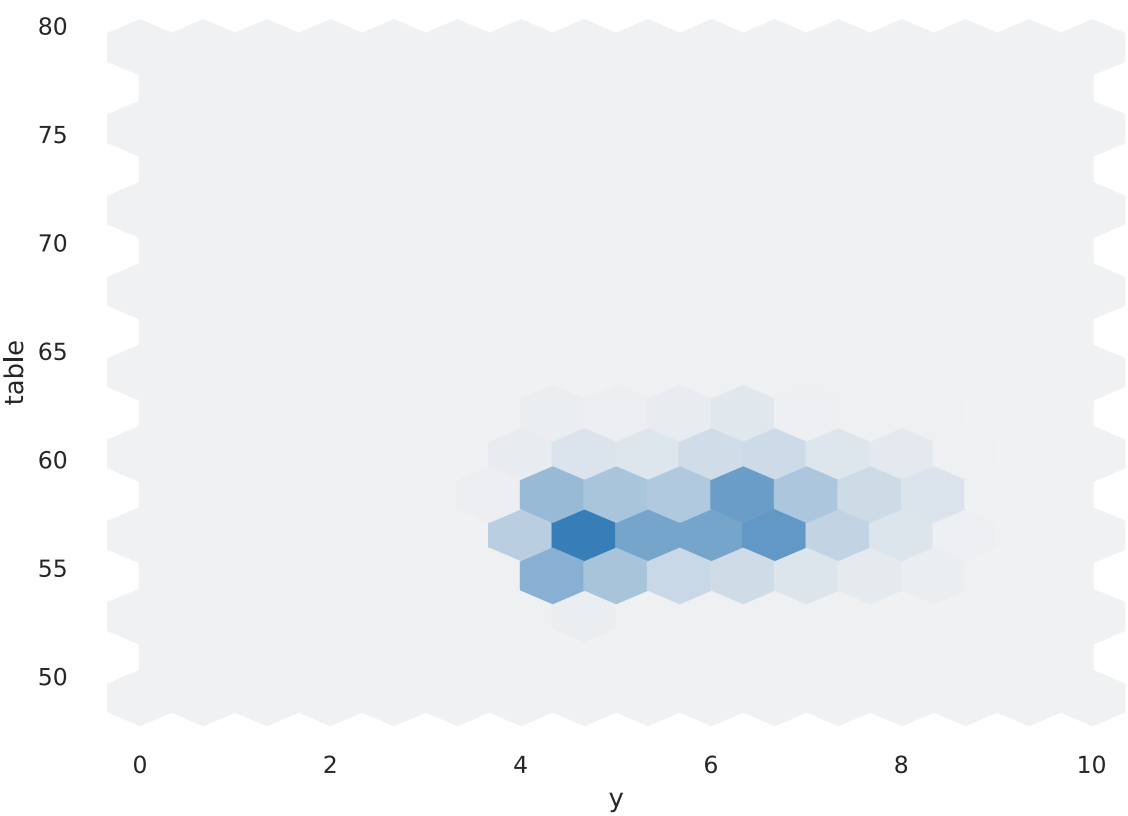


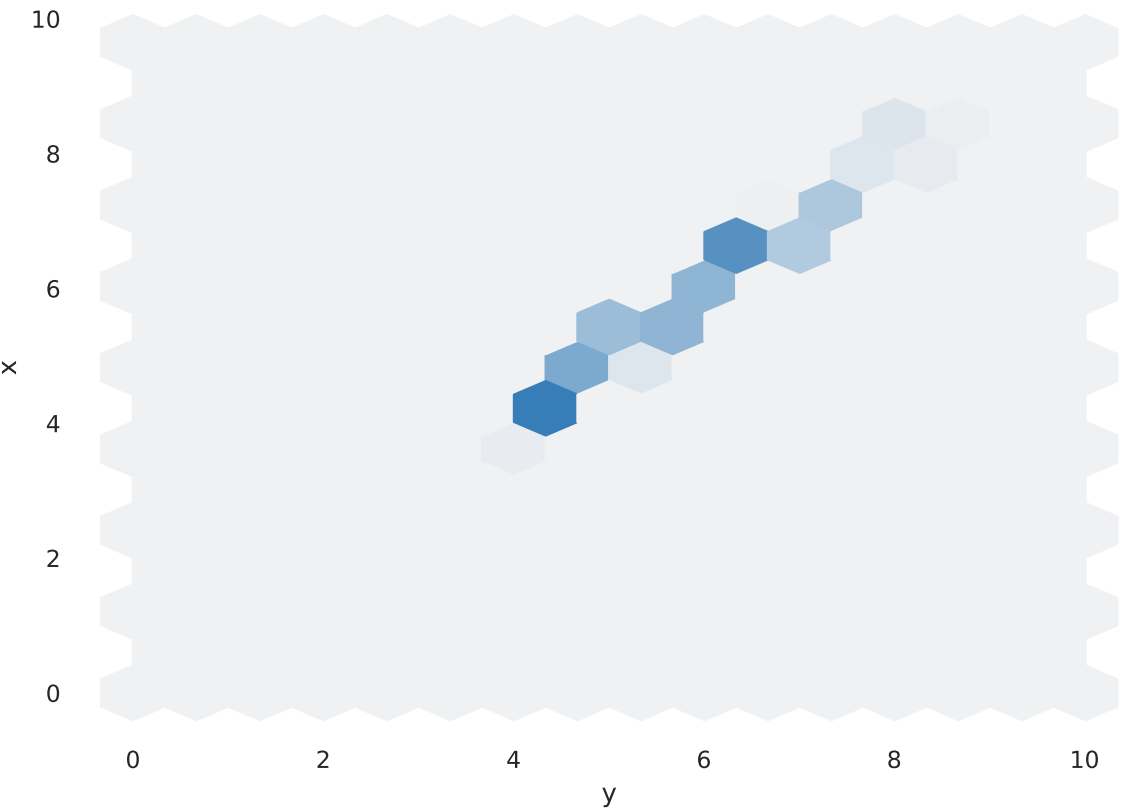


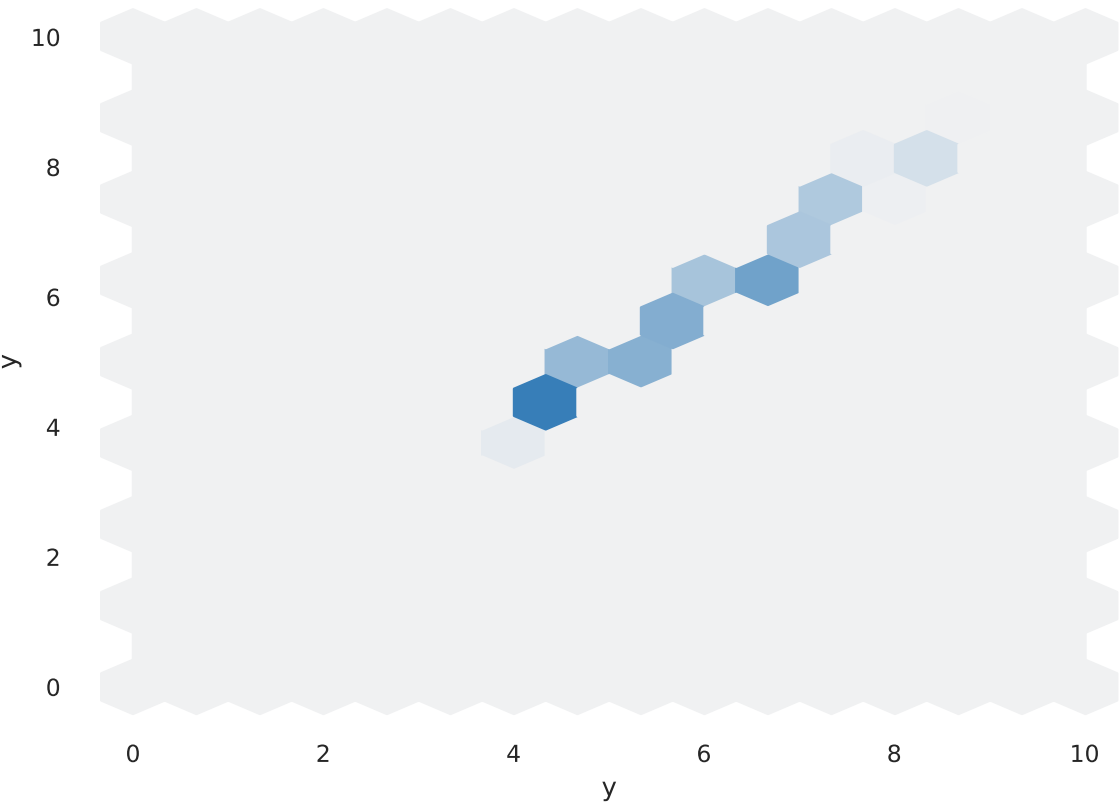


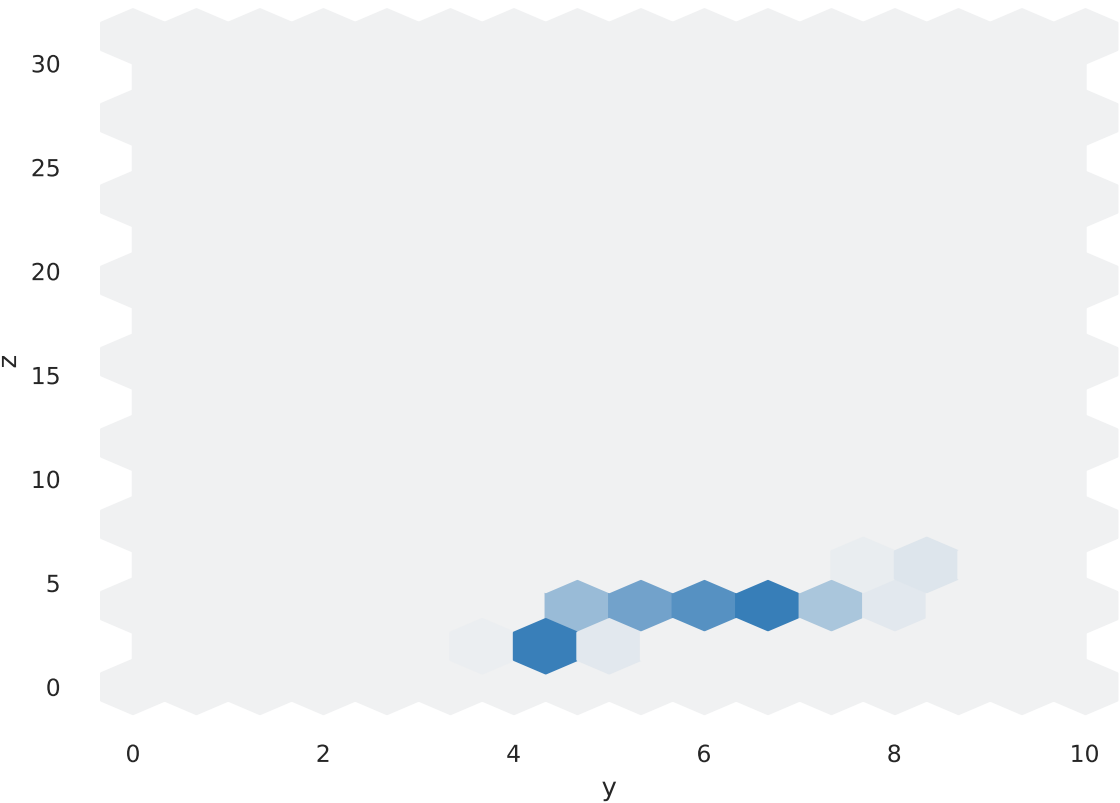


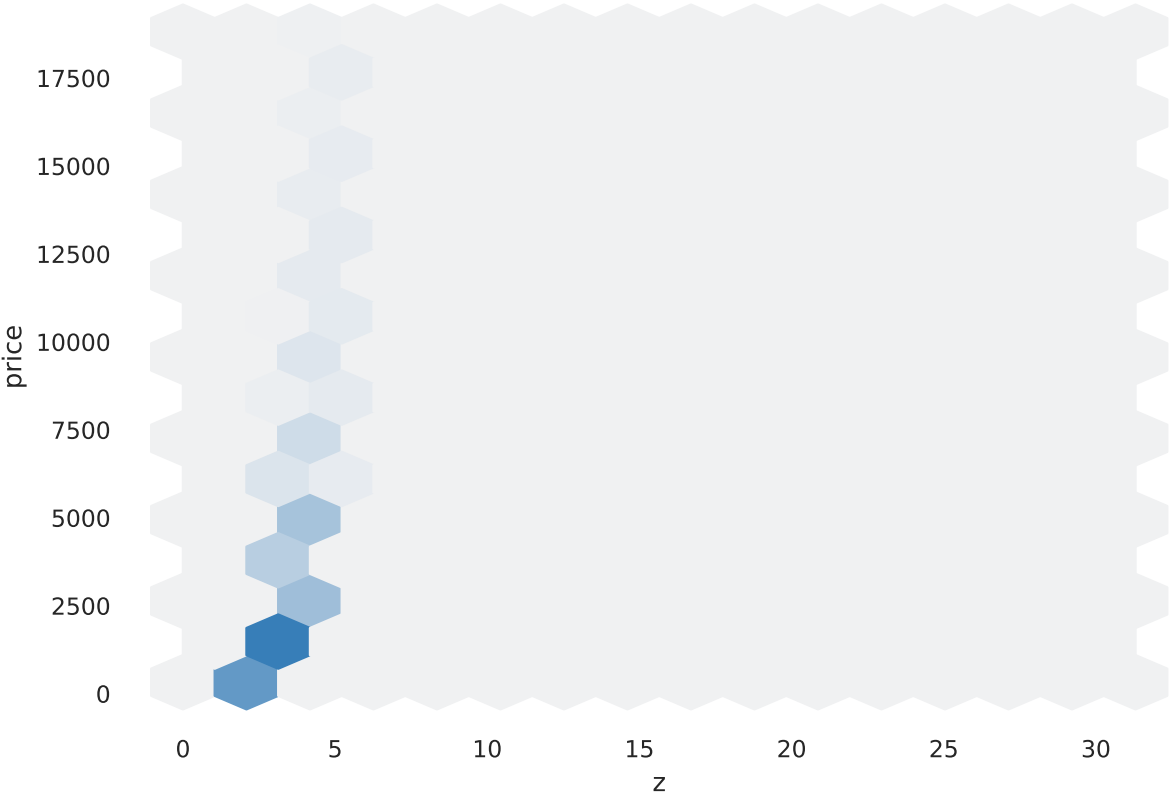


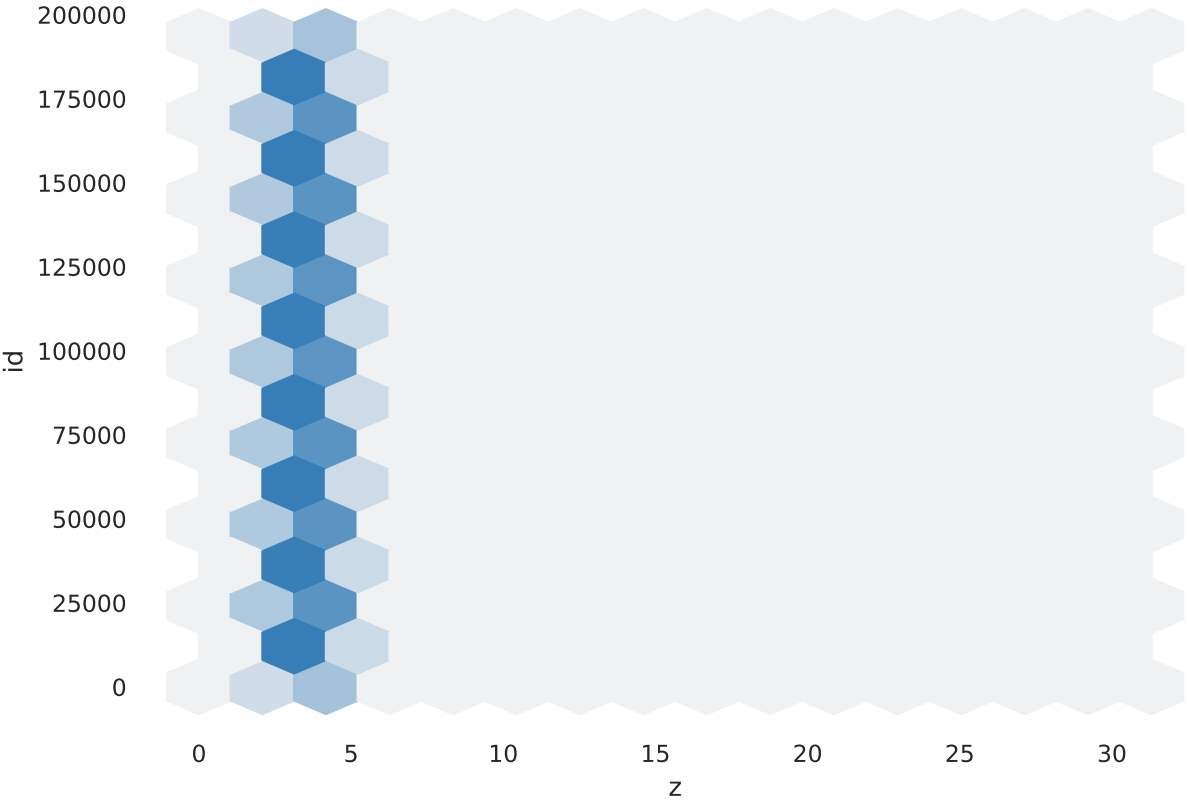


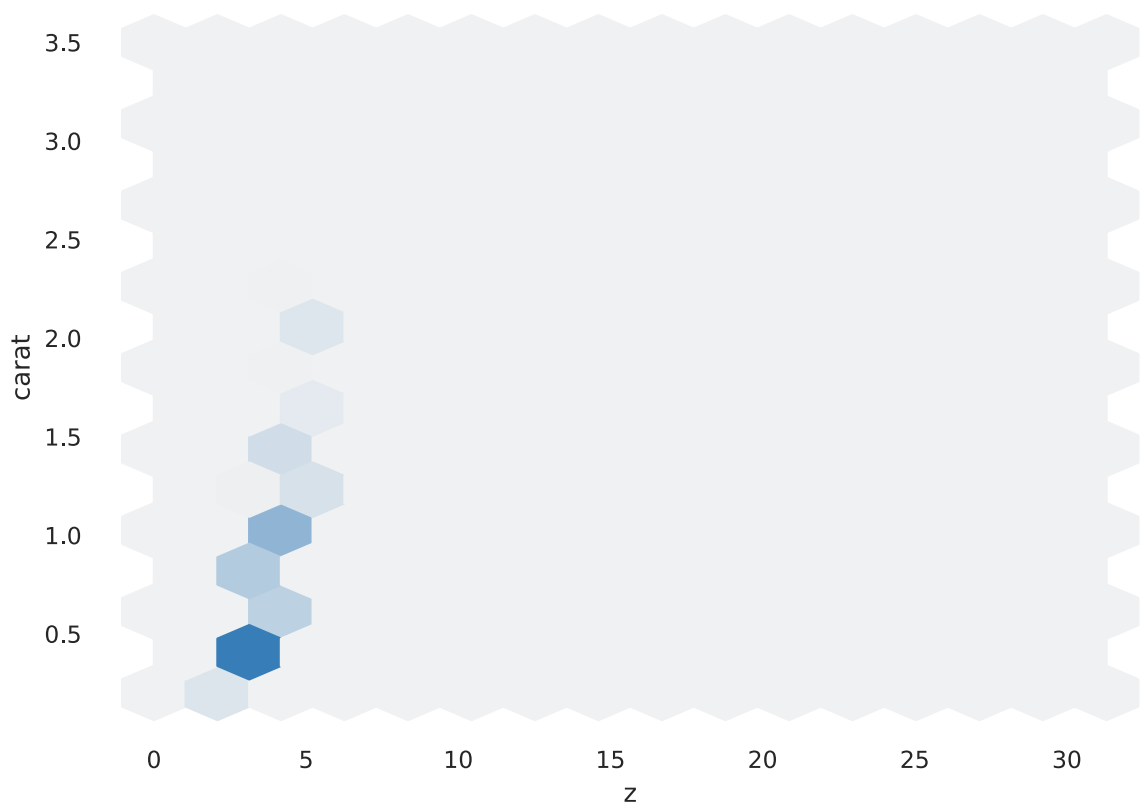


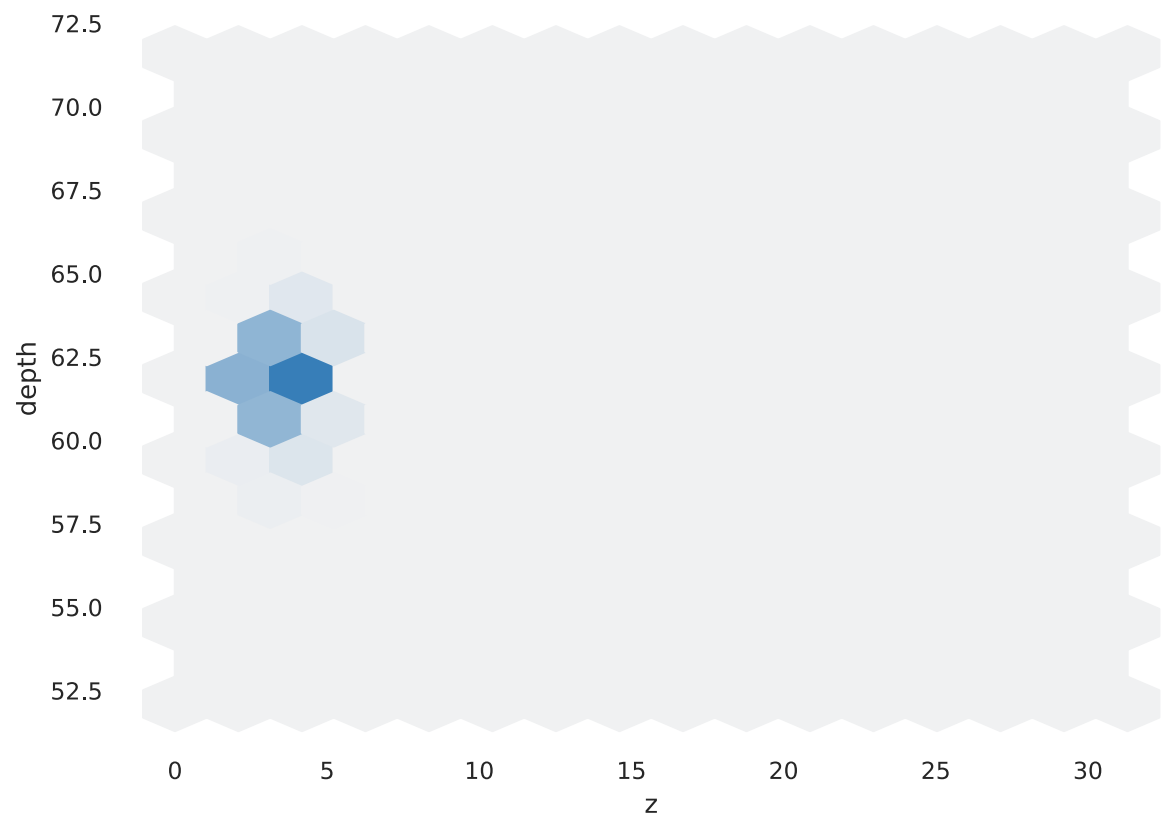


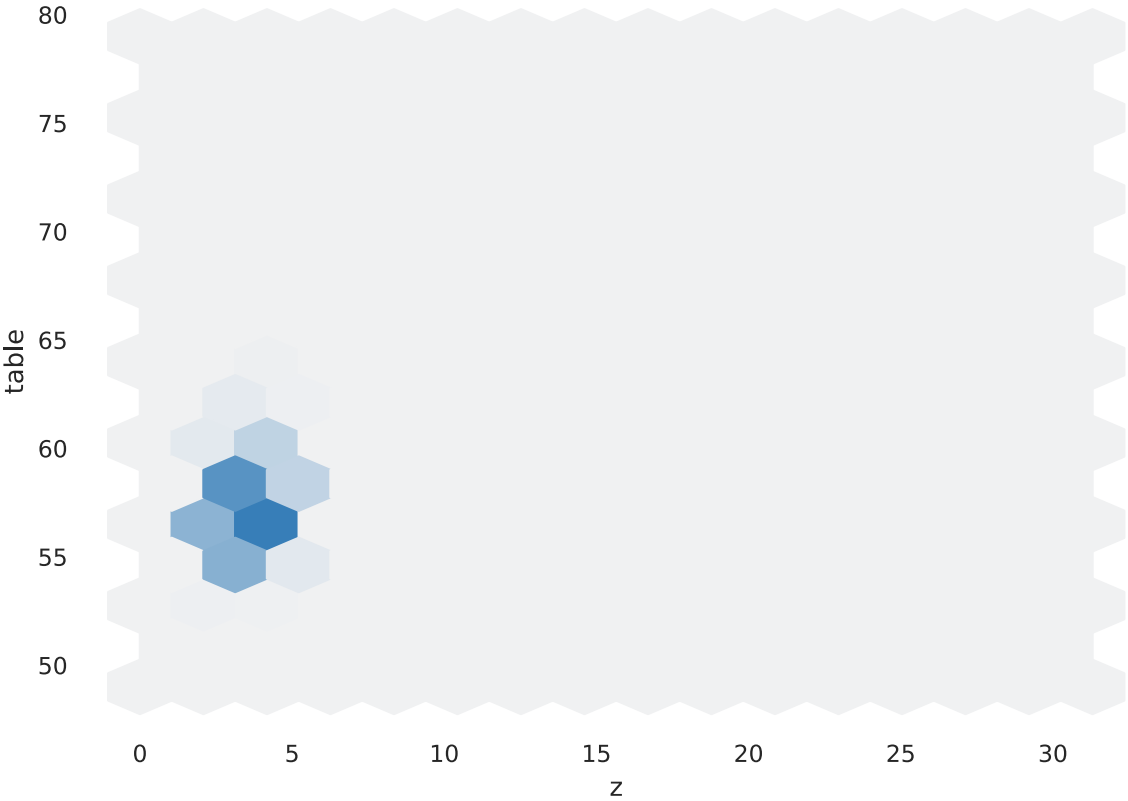


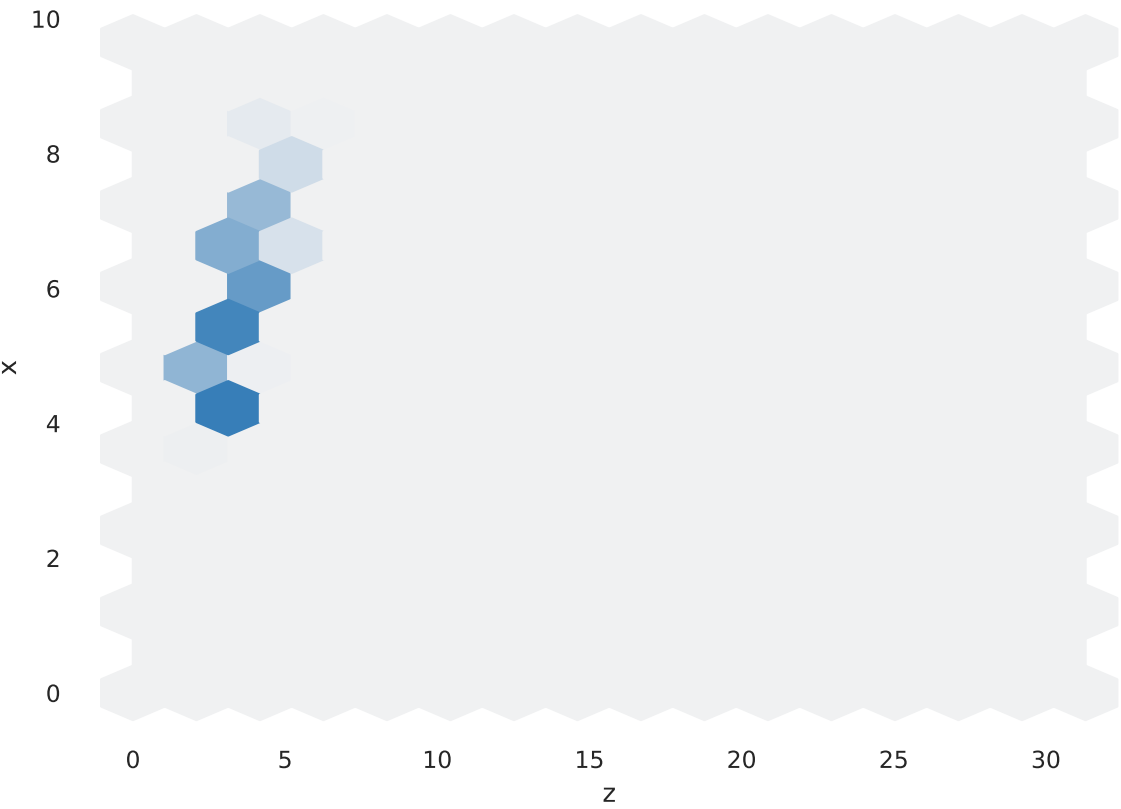


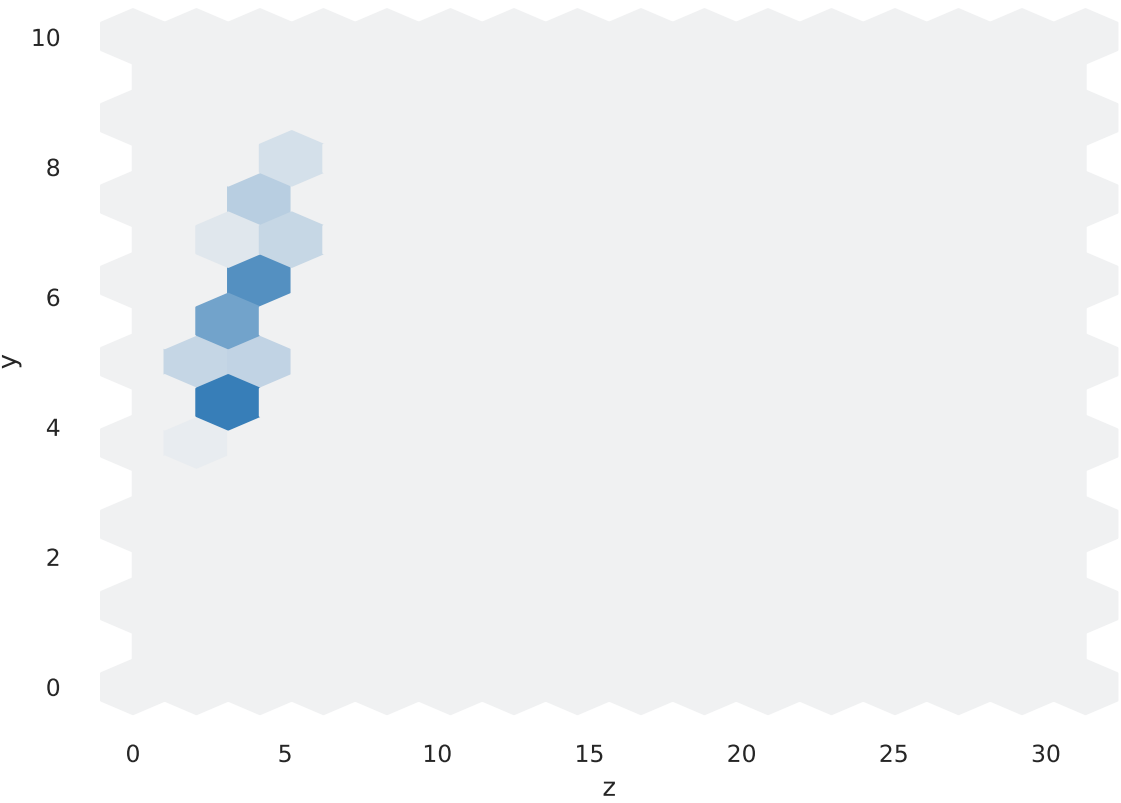


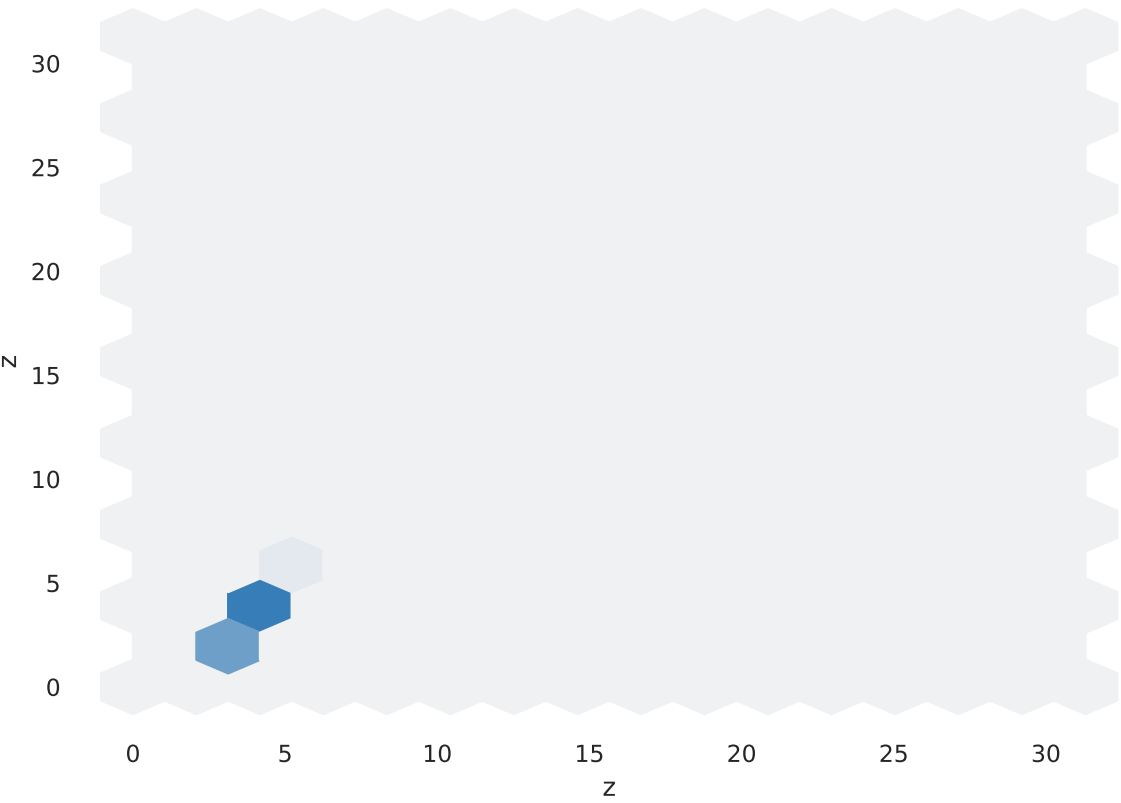


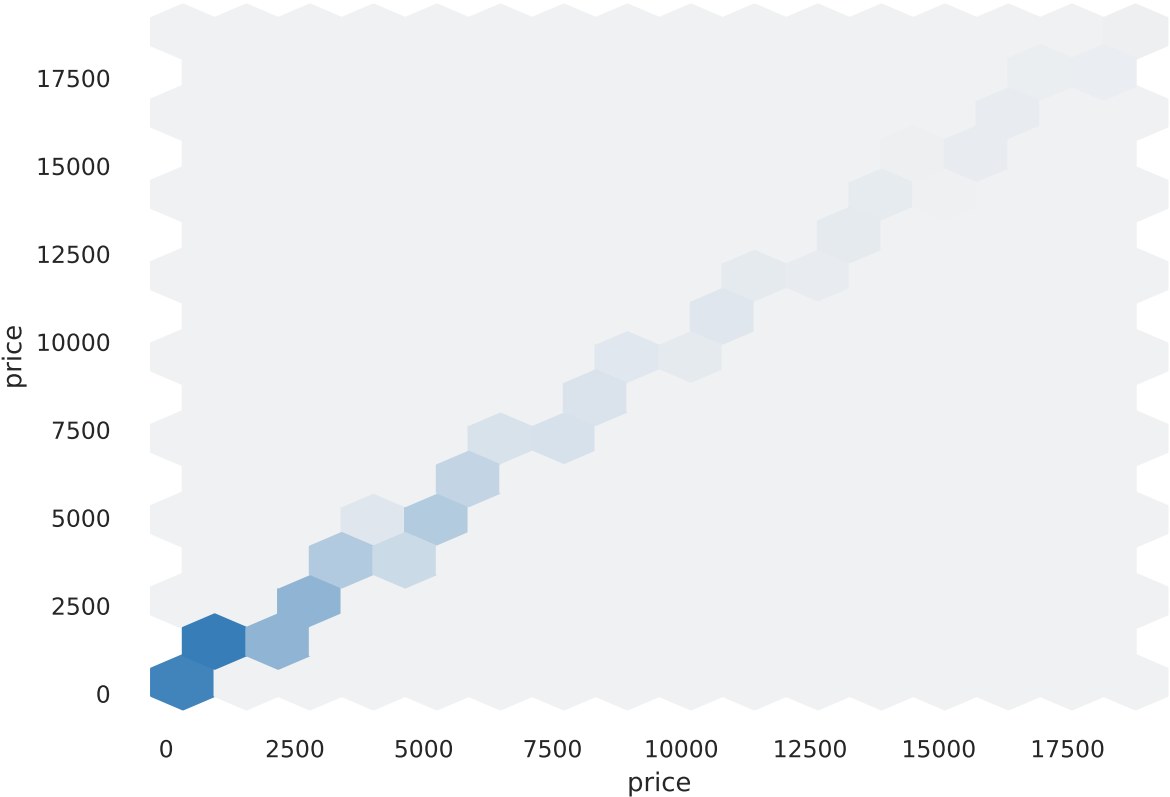


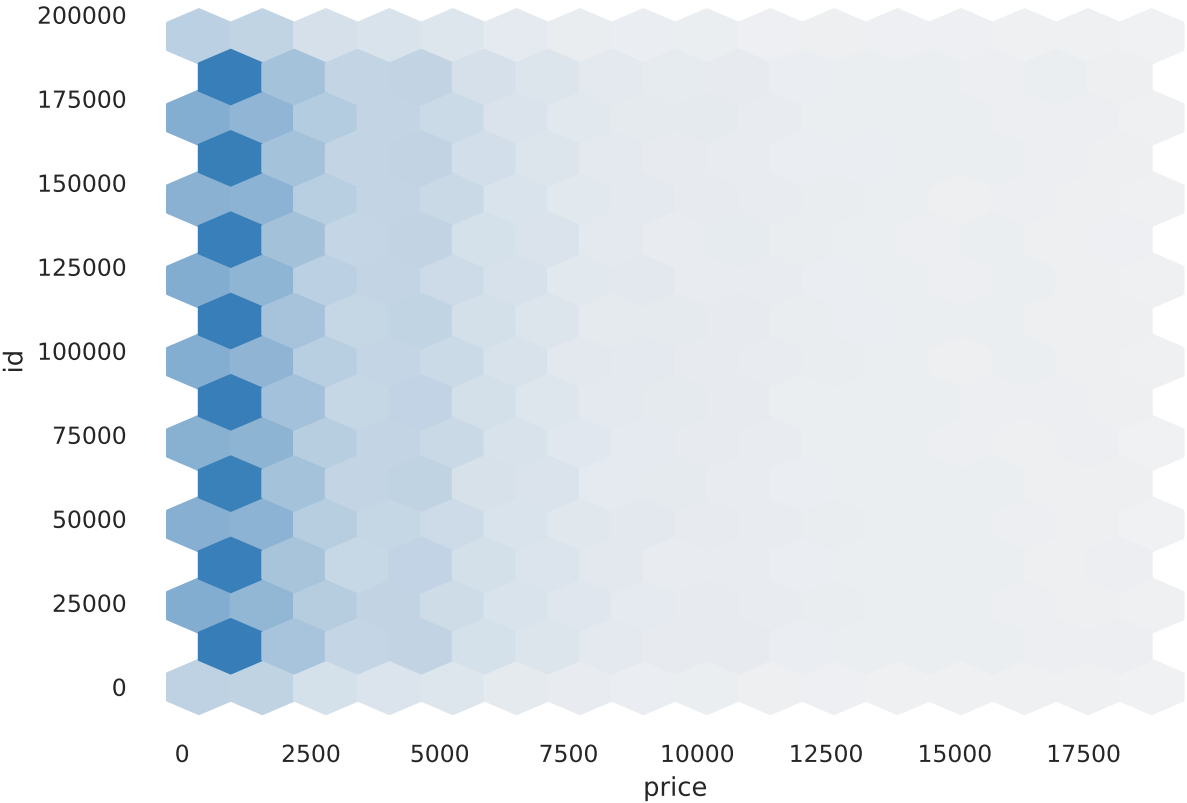


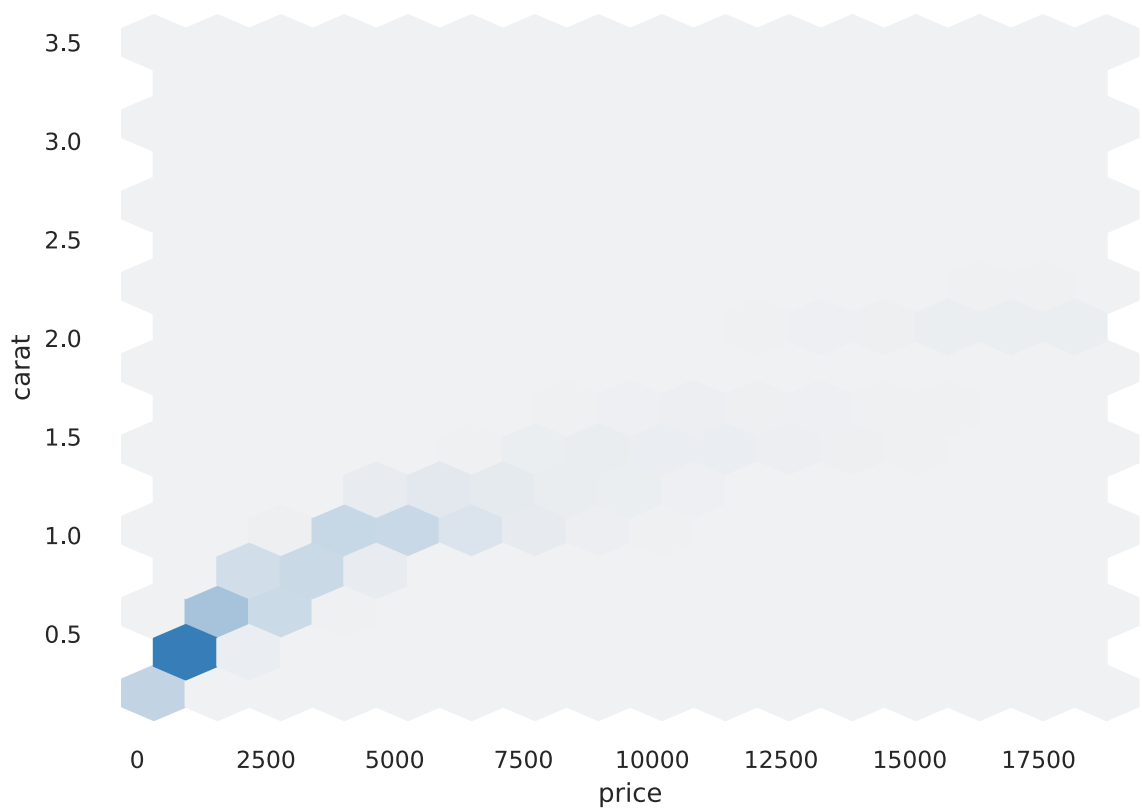


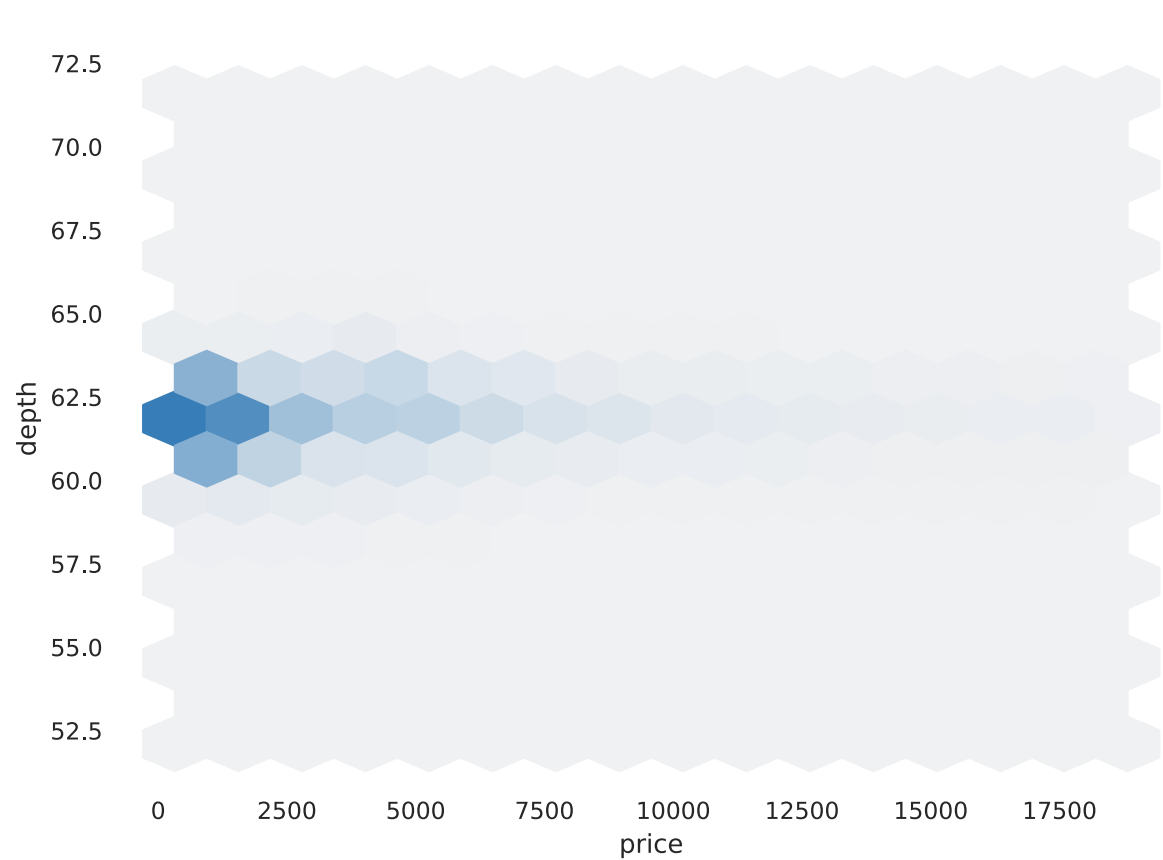


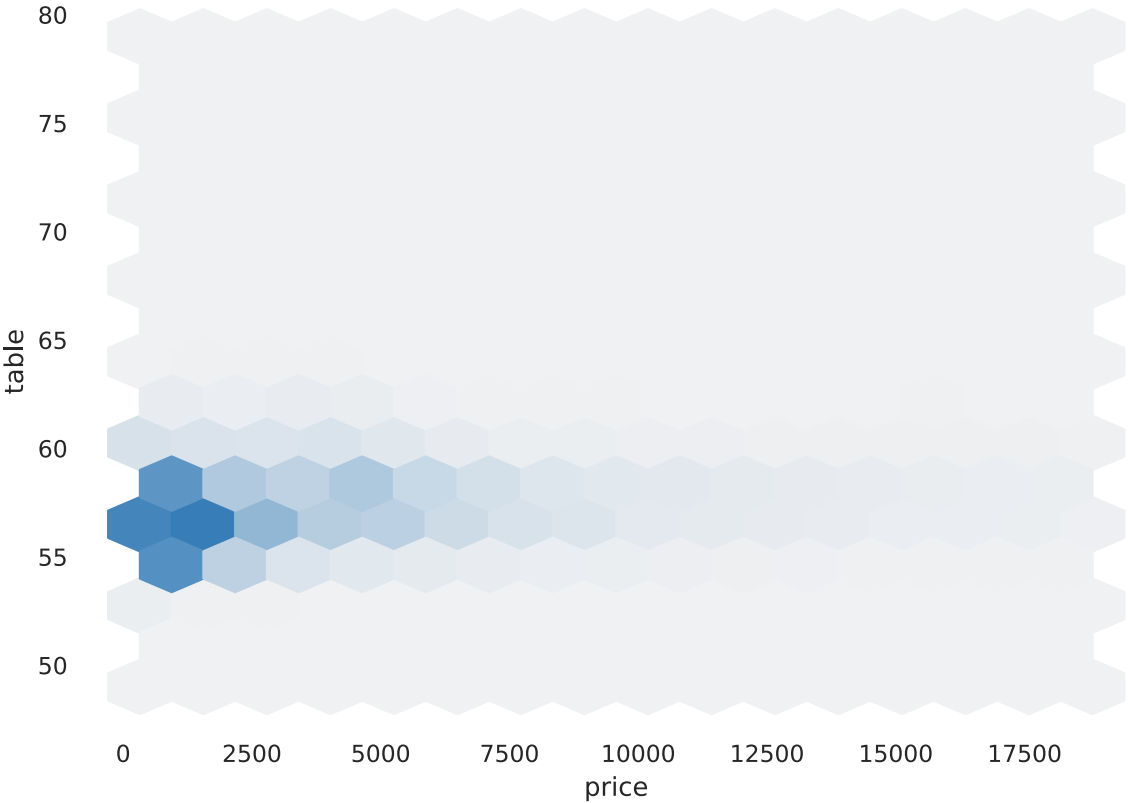


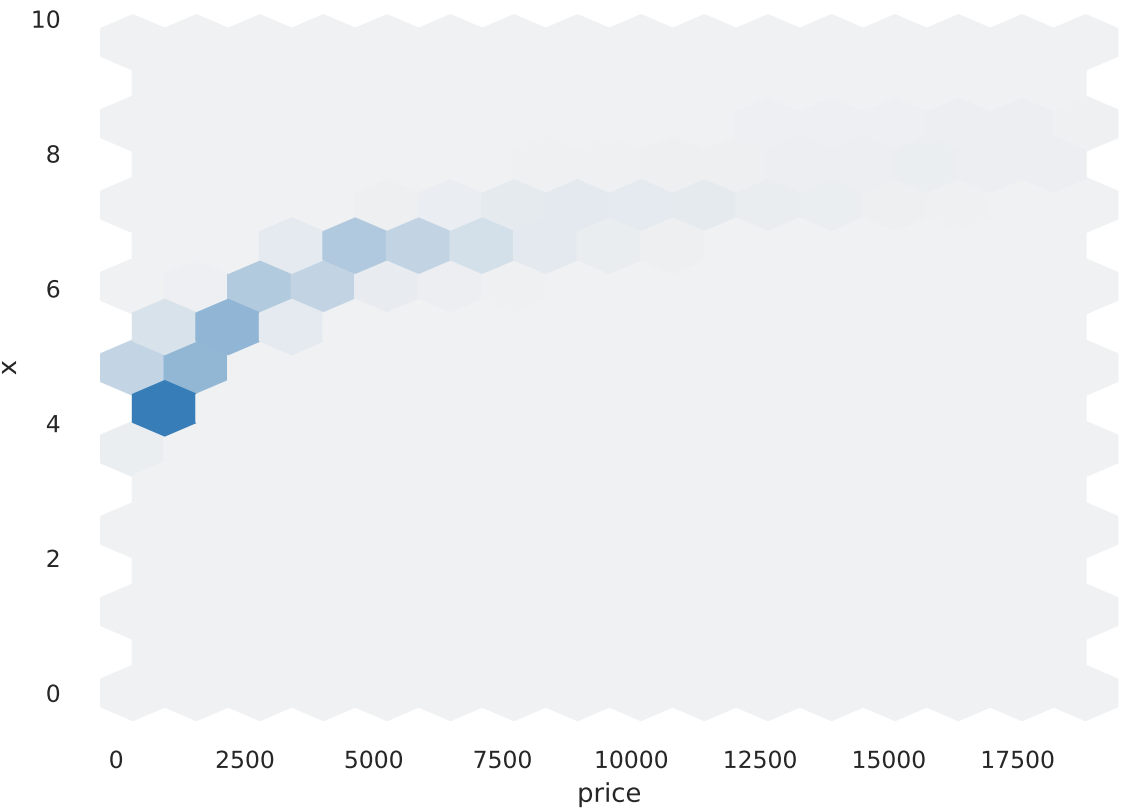


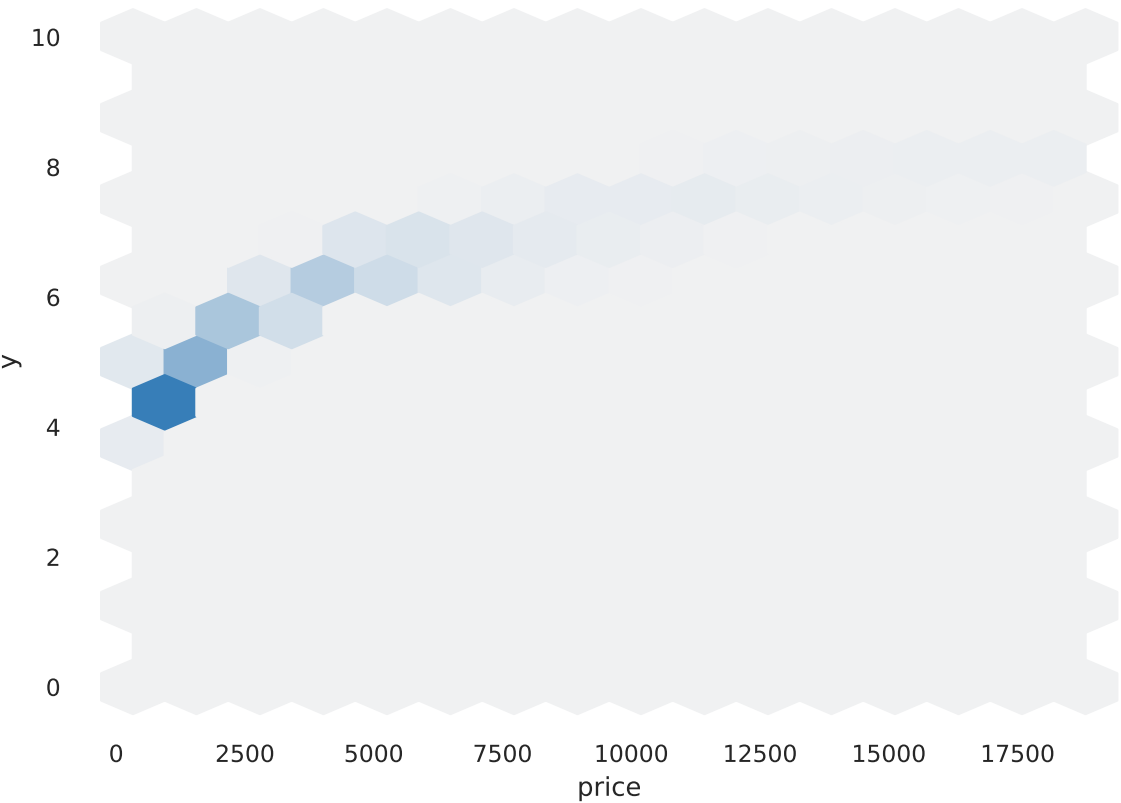


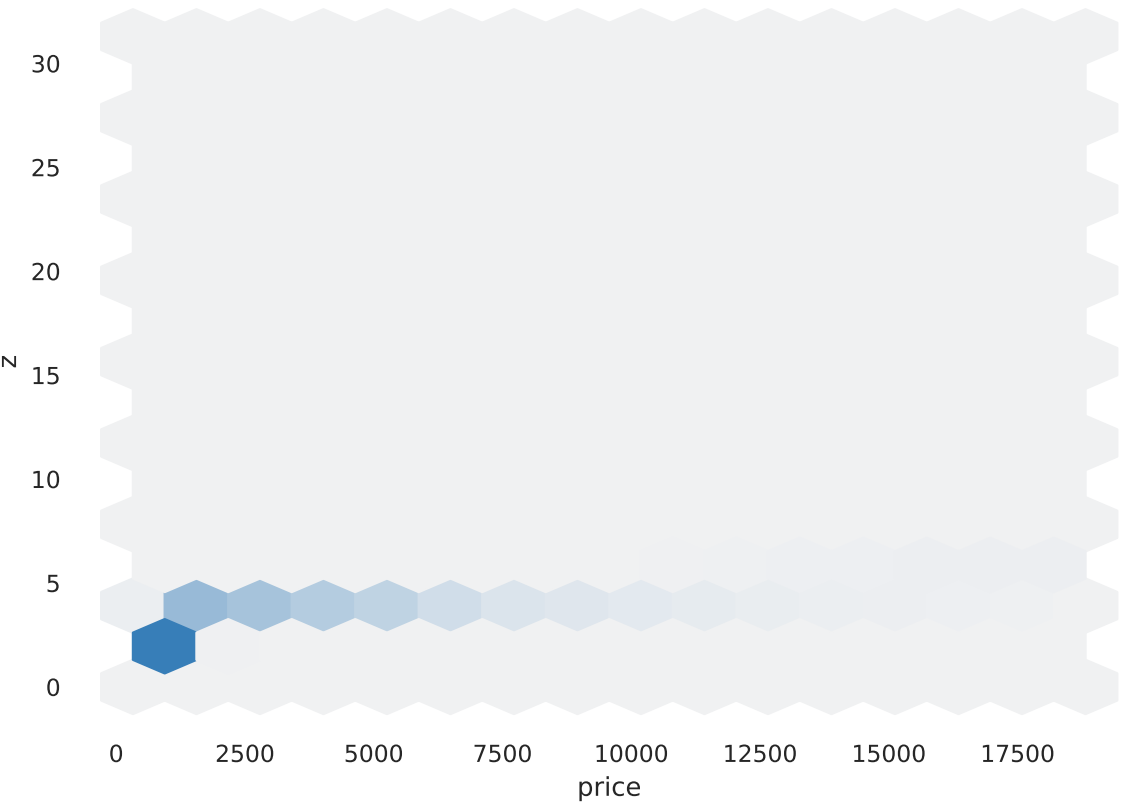




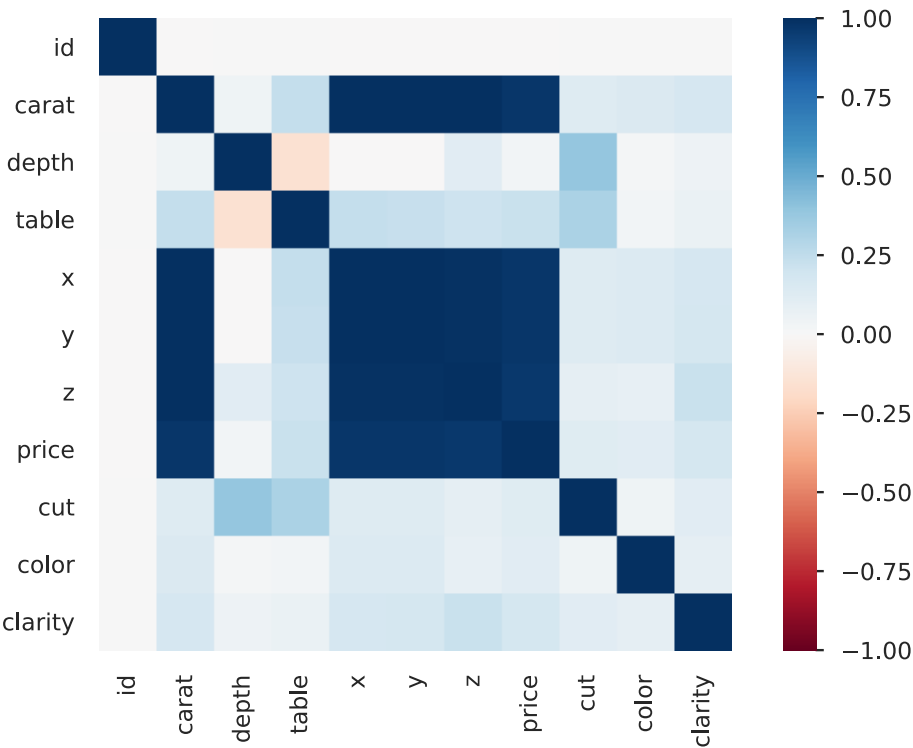






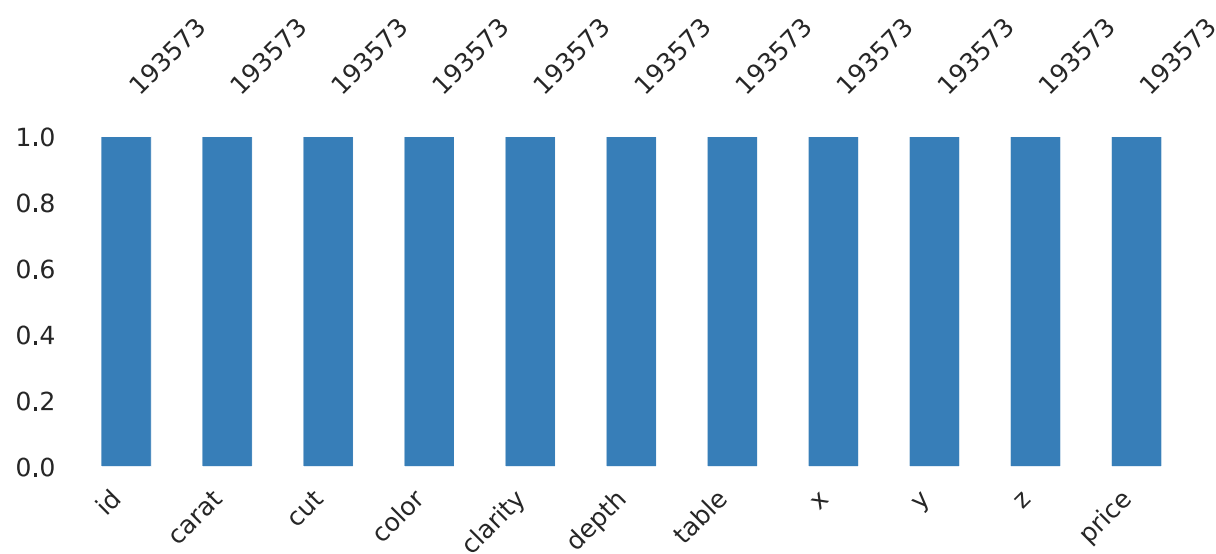


Correlations



	id	carat	depth	table	x	y	z	price	cut	color	clarity
id	1.000	-0.000	0.002	0.002	-0.000	-0.000	-0.000	-0.000	0.001	0.000	0.000
carat	-0.000	1.000	0.042	0.236	0.997	0.996	0.995	0.971	0.131	0.143	0.166
depth	0.002	0.042	1.000	-0.155	-0.007	-0.008	0.110	0.030	0.390	0.022	0.052
table	0.002	0.236	-0.155	1.000	0.240	0.234	0.211	0.223	0.318	0.027	0.063
x	-0.000	0.997	-0.007	0.240	1.000	0.998	0.990	0.972	0.128	0.140	0.170
y	-0.000	0.996	-0.008	0.234	0.998	1.000	0.990	0.972	0.131	0.137	0.172
z	-0.000	0.995	0.110	0.211	0.990	0.990	1.000	0.968	0.093	0.081	0.219
price	-0.000	0.971	0.030	0.223	0.972	0.972	0.968	1.000	0.119	0.116	0.172
cut	0.001	0.131	0.390	0.318	0.128	0.131	0.093	0.119	1.000	0.041	0.115
color	0.000	0.143	0.022	0.027	0.140	0.137	0.081	0.116	0.041	1.000	0.088

Missing values



A simple visualization of nullity by column.

	id	carat	cut	color	clarity	depth	table	x	y	z	price
1											
193573											

Nullity matrix is a data-dense display which lets you quickly visually pick out patterns in data completion.

Sample

id		carat	cut	color	clarity	depth	table	x	y	z	price
0	0	1.52	Premium	F	VS2	62.2	58.0	7.27	7.33	4.55	13619
1	1	2.03	Very Good	J	SI2	62.0	58.0	8.06	8.12	5.05	13387
2	2	0.70	Ideal	G	VS1	61.2	57.0	5.69	5.73	3.50	2772
3	3	0.32	Ideal	G	VS1	61.6	56.0	4.38	4.41	2.71	666
4	4	1.70	Premium	G	VS2	62.6	59.0	7.65	7.61	4.77	14453
5	5	1.51	Very Good	J	SI1	62.8	58.0	7.34	7.29	4.59	7506
6	6	0.74	Ideal	E	VS2	61.8	57.0	5.76	5.79	3.57	3229
7	7	1.34	Premium	G	SI2	62.5	57.0	7.00	7.05	4.38	6224
8	8	0.30	Ideal	F	IF	62.0	56.0	4.35	4.37	2.70	886
9	9	0.30	Good	J	VS1	63.6	57.0	4.26	4.28	2.72	421

id		carat	cut	color	clarity	depth	table	x	y	z	price
193563	193563	0.28	Very Good	E	SI1	61.2	55.0	4.23	4.26	2.60	484
193564	193564	0.90	Good	D	SI1	63.2	57.0	6.11	6.14	3.88	4919
193565	193565	0.31	Premium	D	SI1	62.6	58.0	4.32	4.29	2.69	732
193566	193566	1.05	Ideal	G	VS1	62.1	56.0	6.54	6.51	4.06	7397
193567	193567	0.58	Ideal	E	VS2	61.8	57.0	5.33	5.36	3.31	1872
193568	193568	0.31	Ideal	D	VVS2	61.1	56.0	4.35	4.39	2.67	1130
193569	193569	0.70	Premium	G	VVS2	60.3	58.0	5.75	5.77	3.47	2874
193570	193570	0.73	Very Good	F	SI1	63.1	57.0	5.72	5.75	3.62	3036
193571	193571	0.34	Very Good	D	SI1	62.9	55.0	4.45	4.49	2.81	681
193572	193572	0.71	Good	E	SI2	60.8	64.0	5.73	5.71	3.48	2258



Report generated by YData (https://ydata.ai/?utm_source=opensource&utm_medium=pandasprofiling&utm_campaign=report).