

Overview

Dataset statistics		Variable types	
Number of variables	7	Categorical	2
Number of observations	10000	Numeric	5
Missing cells	0		
Missing cells (%)	0.0%		
Duplicate rows	0		
Duplicate rows (%)	0.0%		
Total size in memory	547.0 KiB		
Average record size in memory	56.0 B		

Alerts

Air temperature [K]	is highly overall correlated with	Process temperature [K]	High correlation
Process temperature [K]	is highly overall correlated with	Air temperature [K]	High correlation
Rotational speed [rpm]	is highly overall correlated with	Torque [Nm]	High correlation
Torque [Nm]	is highly overall correlated with	Rotational speed [rpm]	High correlation
Machine failure	is highly imbalanced (78.6%)		Imbalance
Tool wear [min]	has 120 (1.2%) zeros		Zeros

Reproduction

Analysis started	2023-09-27 09:16:11.490438
Analysis finished	2023-09-27 09:16:21.415705
Duration	9.93 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%22Pandas%20Profiling%20Report%22%2C%20%22dataset%22%3A%20%7B%22description%22%3A%20%22This%20report%20was%20generated%20by%20the%20Pandas%20Profiling%20library%22%2C%20%22author%22%3A%20%22DataCamp%22%2C%20%22version%22%3A%20%223.6.6%22%7D%7D)

Variables

Select Columns

Type
Categorical

Distinct	3
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	78.2 KiB

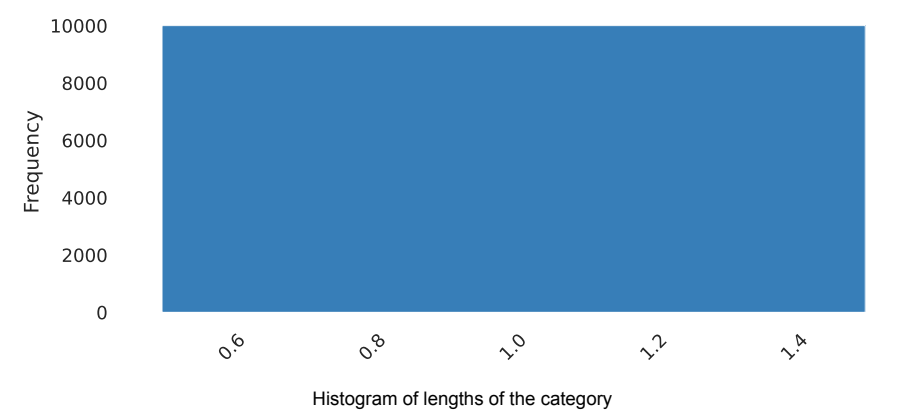
Length		Characters and Unicode	Unique		Sample	
Max length	1		Unique	0 ?	1st row	M
Median length	1		Unique (%)	0.0%	2nd row	L
Mean length	1				3rd row	L

Min length	1	Distinct characters	3	4th row	L
		Distinct categories	1 (https://en.wikipedia.org/wiki/Unicode_character_property#General_Category)	5th row	L
		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)		?
<p>The Unicode Standard assigns character properties to each code point, which can be used to analyse textual variables.</p>					

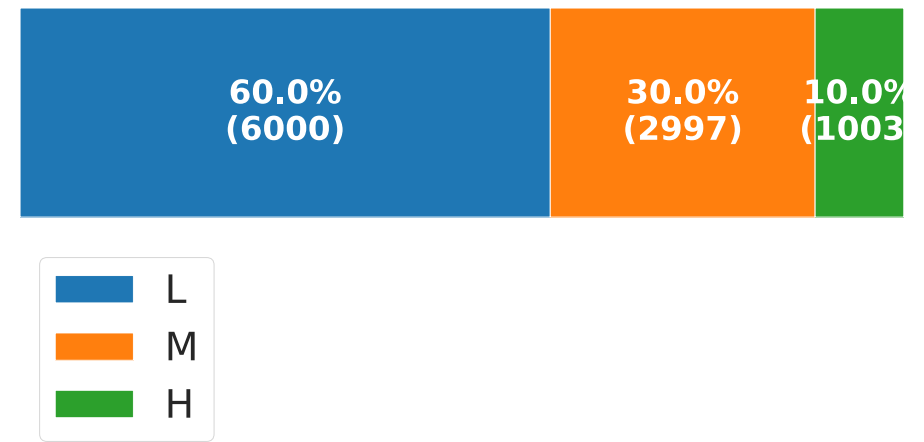
Common Values

Value	Count	Frequency (%)
L	6000	60.0%
M	2997	30.0%
H	1003	10.0%

Length



Common Values (Plot)



Value	Count	Frequency (%)
l	6000	60.0%
m	2997	30.0%
h	1003	10.0%

Most occurring characters

Value	Count	Frequency (%)
L	6000	60.0%

Value	Count	Frequency (%)
M	2997	30.0%
H	1003	10.0%

Most occurring categories

Value	Count	Frequency (%)
Uppercase Letter	10000	100.0%

Most frequent character per category

Uppercase Letter

Value	Count	Frequency (%)
L	6000	60.0%
M	2997	30.0%
H	1003	10.0%

Most occurring scripts

Value	Count	Frequency (%)
Latin	10000	100.0%

Most frequent character per script

Latin

Value	Count	Frequency (%)
L	6000	60.0%
M	2997	30.0%
H	1003	10.0%

Most occurring blocks

Value	Count	Frequency (%)
ASCII	10000	100.0%

Most frequent character per block

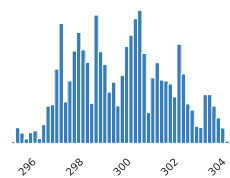
ASCII

Value	Count	Frequency (%)
L	6000	60.0%
M	2997	30.0%
H	1003	10.0%

Air temperature [K]
Real number (\mathbb{R})

Distinct	93
Distinct (%)	0.9%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Mean	300.00493

Minimum	295.3
Maximum	304.5
Zeros	0
Zeros (%)	0.0%
Negative	0
Negative (%)	0.0%
Memory size	78.2 KiB

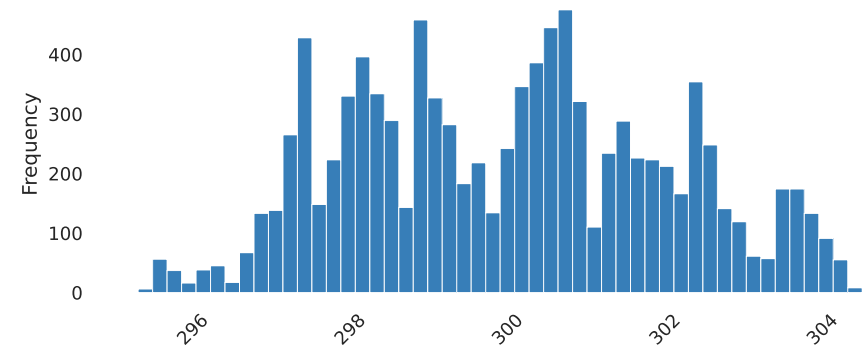


Quantile statistics

Minimum	295.3
5-th percentile	297.1
Q1	298.3
median	300.1
Q3	301.5
95-th percentile	303.5
Maximum	304.5
Range	9.2
Interquartile range (IQR)	3.2

Descriptive statistics

Standard deviation	2.0002587
Coefficient of variation (CV)	0.0066674194
Kurtosis	-0.83596167
Mean	300.00493
Median Absolute Deviation (MAD)	1.6
Skewness	0.11427392
Sum	3000049.3
Variance	4.0010348
Monotonicity	Not monotonic



Histogram with fixed size bins (bins=50)

Value	Count	Frequency (%)
300.7	279	2.8%
298.9	231	2.3%
297.4	230	2.3%
300.5	229	2.3%
298.8	227	2.3%
300.6	216	2.2%
298.2	208	2.1%
302.3	203	2.0%
297.5	198	2.0%
300.4	198	2.0%
Other values (83)	7781	77.8%

Value	Count	Frequency (%)
295.3	3	< 0.1%
295.4	3	< 0.1%

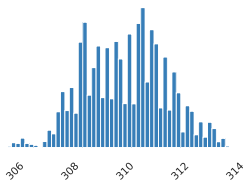
Value	Count	Frequency (%)
295.5	18	0.2%
295.6	38	0.4%
295.7	18	0.2%
295.8	19	0.2%
295.9	10	0.1%
296	6	0.1%
296.1	12	0.1%
296.2	26	0.3%

Value	Count	Frequency (%)
304.5	1	< 0.1%
304.4	7	0.1%
304.3	15	0.1%
304.2	40	0.4%
304.1	46	0.5%
304	45	0.4%
303.9	58	0.6%
303.8	75	0.8%
303.7	96	1.0%
303.6	78	0.8%

Process temperature [K]
Real number (\mathbb{R})

Distinct	82
Distinct (%)	0.8%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Mean	310.00556

Minimum	305.7
Maximum	313.8
Zeros	0
Zeros (%)	0.0%
Negative	0
Negative (%)	0.0%
Memory size	78.2 KiB

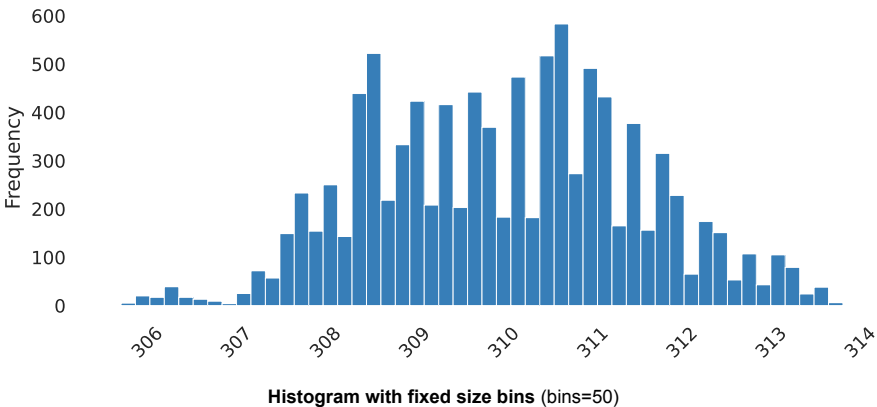


Quantile statistics

Minimum	305.7
5-th percentile	307.7
Q1	308.8
median	310.1
Q3	311.1
95-th percentile	312.5
Maximum	313.8
Range	8.1
Interquartile range (IQR)	2.3

Descriptive statistics

Standard deviation	1.4837342
Coefficient of variation (CV)	0.0047861536
Kurtosis	-0.49973437
Mean	310.00556
Median Absolute Deviation (MAD)	1.1
Skewness	0.015027268
Sum	3100055.6
Variance	2.2014672
Monotonicity	Not monotonic



Value	Count	Frequency (%)
310.6	317	3.2%
310.8	273	2.7%
310.7	266	2.7%
308.6	265	2.6%
310.5	263	2.6%
310.1	260	2.6%
308.5	257	2.6%
310.4	254	2.5%
311	246	2.5%
310.9	245	2.5%
Other values (72)	7354	73.5%

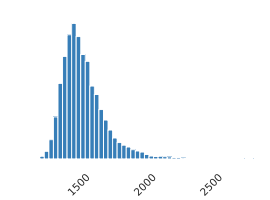
Value	Count	Frequency (%)
305.7	2	< 0.1%
305.8	3	< 0.1%

Value	Count	Frequency (%)
305.9	6	0.1%
306	14	0.1%
306.1	17	0.2%
306.2	26	0.3%
306.3	13	0.1%
306.4	8	0.1%
306.5	9	0.1%
306.6	13	0.1%

Value	Count	Frequency (%)
313.8	2	< 0.1%
313.7	4	< 0.1%
313.6	16	0.2%
313.5	22	0.2%
313.4	24	0.2%
313.3	29	0.3%
313.2	50	0.5%
313.1	50	0.5%
313	55	0.5%
312.9	43	0.4%

Rotational speed [rpm]
Real number (\mathbb{R})

Distinct	941	Minimum	1168
Distinct (%)	9.4%	Maximum	2886
Missing	0	Zeros	0
Missing (%)	0.0%	Zeros (%)	0.0%
Infinite	0	Negative	0
Infinite (%)	0.0%	Negative (%)	0.0%
Mean	1538.7761	Memory size	78.2 KiB

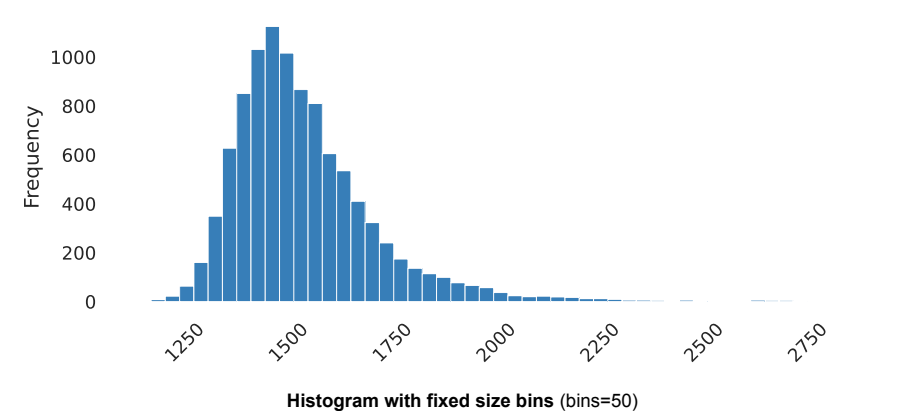


Quantile statistics

Minimum	1168
5-th percentile	1332
Q1	1423
median	1503
Q3	1612
95-th percentile	1868.05
Maximum	2886
Range	1718
Interquartile range (IQR)	189

Descriptive statistics

Standard deviation	179.2841
Coefficient of variation (CV)	0.11651084
Kurtosis	7.3929449
Mean	1538.7761
Median Absolute Deviation (MAD)	91
Skewness	1.993171
Sum	15387761
Variance	32142.787
Monotonicity	Not monotonic



Value	Count	Frequency (%)
1452	48	0.5%
1435	43	0.4%
1447	42	0.4%
1429	40	0.4%
1469	40	0.4%
1479	40	0.4%
1450	39	0.4%
1507	39	0.4%
1418	39	0.4%
1446	38	0.4%
Other values (931)	9592	95.9%

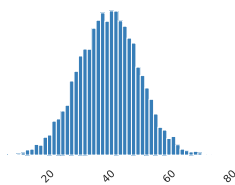
Value	Count	Frequency (%)
1168	1	< 0.1%
1181	1	< 0.1%

Value	Count	Frequency (%)
1183	1	< 0.1%
1192	1	< 0.1%
1200	1	< 0.1%
1202	3	< 0.1%
1207	1	< 0.1%
1208	1	< 0.1%
1212	2	< 0.1%
1217	1	< 0.1%

Value	Count	Frequency (%)
2886	1	< 0.1%
2874	1	< 0.1%
2861	1	< 0.1%
2833	1	< 0.1%
2825	1	< 0.1%
2760	1	< 0.1%
2737	1	< 0.1%
2721	1	< 0.1%
2710	1	< 0.1%
2709	1	< 0.1%

Torque [Nm]
Real number (\mathbb{R})

Distinct	577	Minimum	3.8
Distinct (%)	5.8%	Maximum	76.6
Missing	0	Zeros	0
Missing (%)	0.0%	Zeros (%)	0.0%
Infinite	0	Negative	0
Infinite (%)	0.0%	Negative (%)	0.0%
Mean	39.98691	Memory size	78.2 KiB

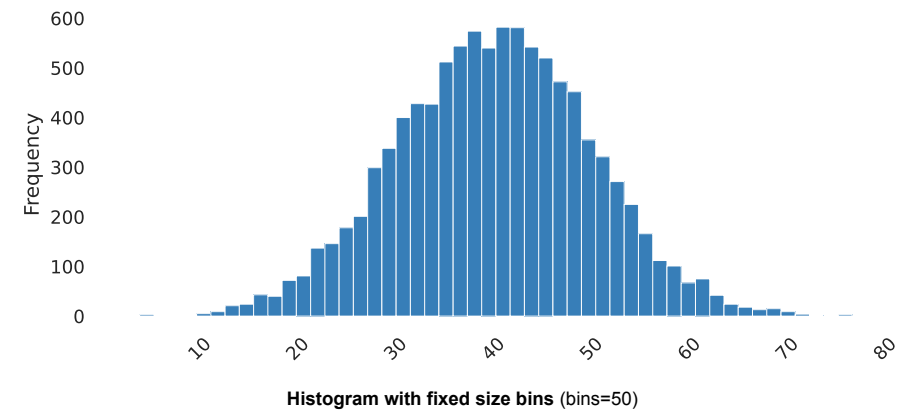


Quantile statistics

Minimum	3.8
5-th percentile	23.5
Q1	33.2
median	40.1
Q3	46.8
95-th percentile	56.1
Maximum	76.6
Range	72.8
Interquartile range (IQR)	13.6

Descriptive statistics

Standard deviation	9.9689337
Coefficient of variation (CV)	0.24930493
Kurtosis	-0.013240614
Mean	39.98691
Median Absolute Deviation (MAD)	6.8
Skewness	-0.0095165958
Sum	399869.1
Variance	99.37964
Monotonicity	Not monotonic



Value	Count	Frequency (%)
40.2	52	0.5%
38.5	50	0.5%
42.4	50	0.5%
35.8	50	0.5%
37.7	49	0.5%
39.9	48	0.5%
40.6	48	0.5%
38.2	48	0.5%
40	47	0.5%
36.6	47	0.5%
Other values (567)	9511	95.1%

Value	Count	Frequency (%)
3.8	1	< 0.1%

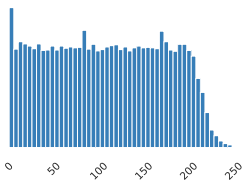
Value	Count	Frequency (%)
4.2	1	< 0.1%
4.6	1	< 0.1%
5.6	1	< 0.1%
5.8	1	< 0.1%
8	1	< 0.1%
8.8	1	< 0.1%
9.3	1	< 0.1%
9.7	2	< 0.1%
9.8	1	< 0.1%

Value	Count	Frequency (%)
76.6	1	< 0.1%
76.2	1	< 0.1%
75.4	1	< 0.1%
74.5	1	< 0.1%
73.6	1	< 0.1%
72.8	1	< 0.1%
72	1	< 0.1%
71.8	1	< 0.1%
71.6	1	< 0.1%
71.3	1	< 0.1%

Tool wear [min]
Real number (\mathbb{R})

Distinct	246
Distinct (%)	2.5%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Mean	107.951

Minimum	0
Maximum	253
Zeros	120
Zeros (%)	1.2%
Negative	0
Negative (%)	0.0%
Memory size	78.2 KiB

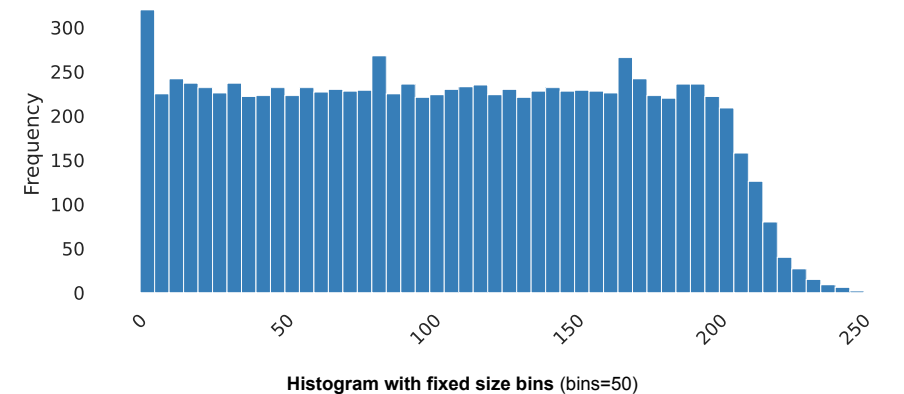


Quantile statistics

Minimum	0
5-th percentile	9.95
Q1	53
median	108
Q3	162
95-th percentile	206.05
Maximum	253
Range	253
Interquartile range (IQR)	109

Descriptive statistics

Standard deviation	63.654147
Coefficient of variation (CV)	0.58965778
Kurtosis	-1.1667371
Mean	107.951
Median Absolute Deviation (MAD)	55
Skewness	0.027292239
Sum	1079510
Variance	4051.8504
Monotonicity	Not monotonic



Value	Count	Frequency (%)
0	120	1.2%
2	69	0.7%
5	63	0.6%
7	58	0.6%
59	58	0.6%
166	57	0.6%
119	57	0.6%
9	55	0.5%
146	54	0.5%
96	54	0.5%
Other values (236)	9355	93.5%

Value	Count	Frequency (%)
0	120	1.2%
2	69	0.7%

Value	Count	Frequency (%)
3	34	0.3%
4	34	0.3%
5	63	0.6%
6	31	0.3%
7	58	0.6%
8	36	0.4%
9	55	0.5%
10	45	0.4%

Value	Count	Frequency (%)
253	1	< 0.1%
251	1	< 0.1%
246	3	< 0.1%
244	3	< 0.1%
242	2	< 0.1%
241	1	< 0.1%
240	3	< 0.1%
239	1	< 0.1%
238	2	< 0.1%
237	1	< 0.1%

Machine failure

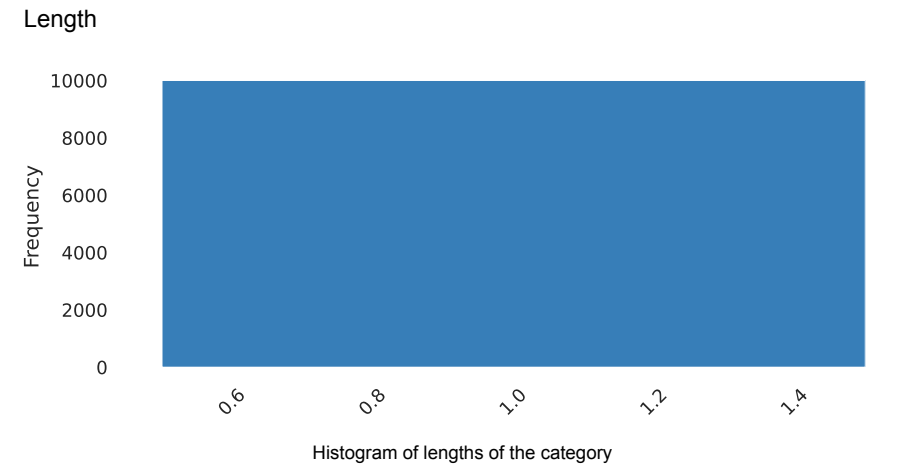
Categorical

Distinct	2
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	78.2 KiB

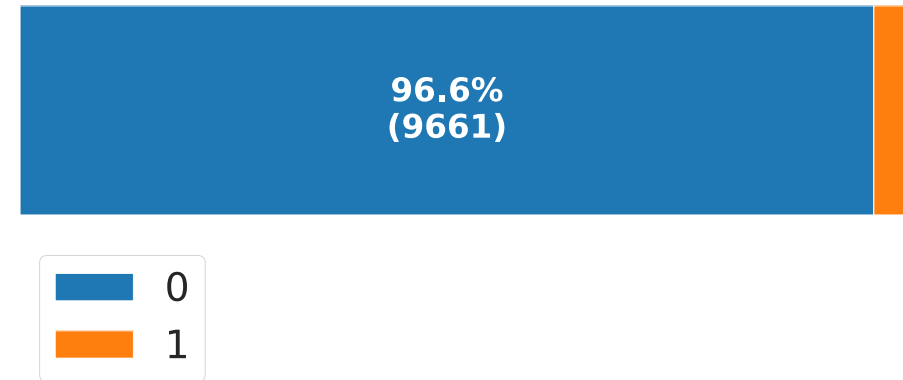
Length		Characters and Unicode		Unique		Sample	
Max length	1	Total characters	10000	Unique	0 ?	1st row	0
Median length	1			Unique (%)	0.0%	2nd row	0
Mean length	1					3rd row	0
Min length	1					4th row	0
						5th row	0 ?
		Distinct categories	1			(https://en.wikipedia.org/wiki/Unicode_character_property#General_Category)	
		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.							

Common Values

Value	Count	Frequency (%)
0	9661	96.6%
1	339	3.4%



Common Values (Plot)



Value	Count	Frequency (%)
0	9661	96.6%
1	339	3.4%

Most occurring characters

Value	Count	Frequency (%)
0	9661	96.6%
1	339	3.4%

Most occurring categories

Value	Count	Frequency (%)
Decimal Number	10000	100.0%

Most frequent character per category

Decimal Number

Value	Count	Frequency (%)
0	9661	96.6%
1	339	3.4%

Most occurring scripts

Value	Count	Frequency (%)
Common	10000	100.0%

Most frequent character per script

Common

Value	Count	Frequency (%)
0	9661	96.6%
1	339	3.4%

Most occurring blocks

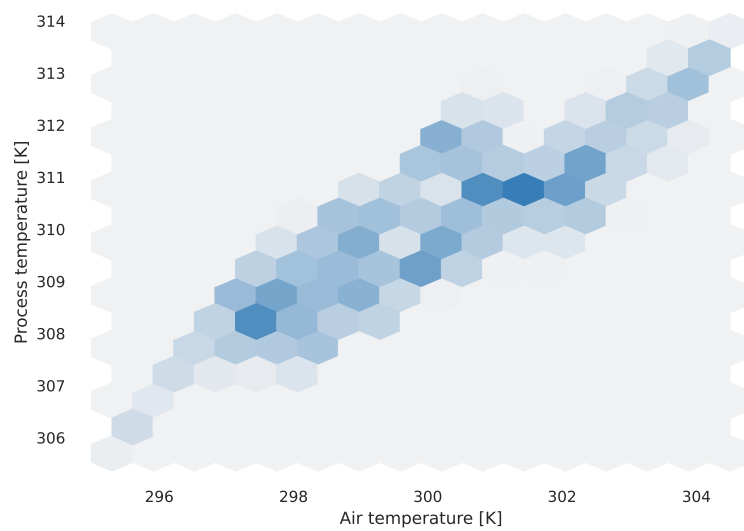
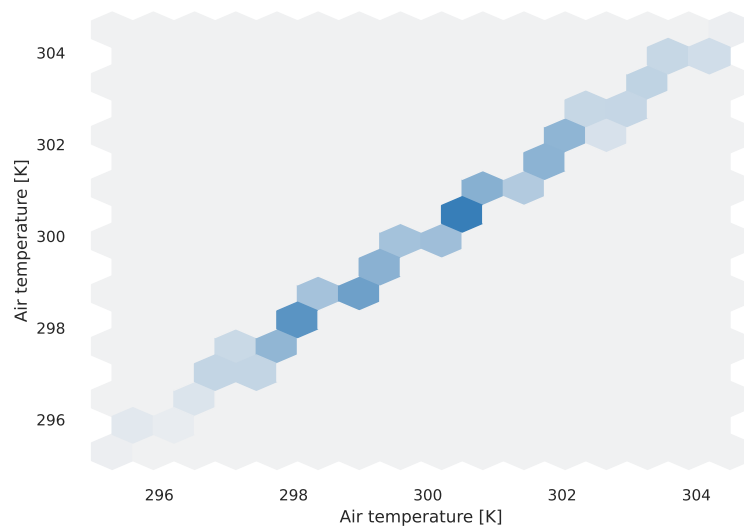
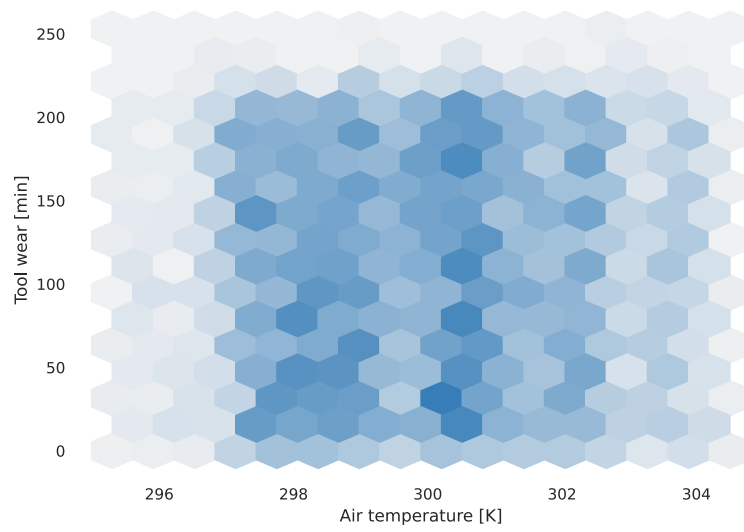
Value	Count	Frequency (%)
ASCII	10000	100.0%

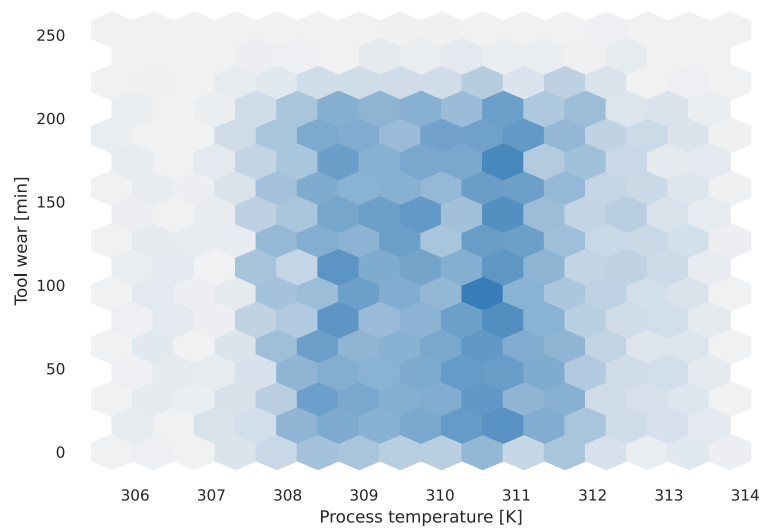
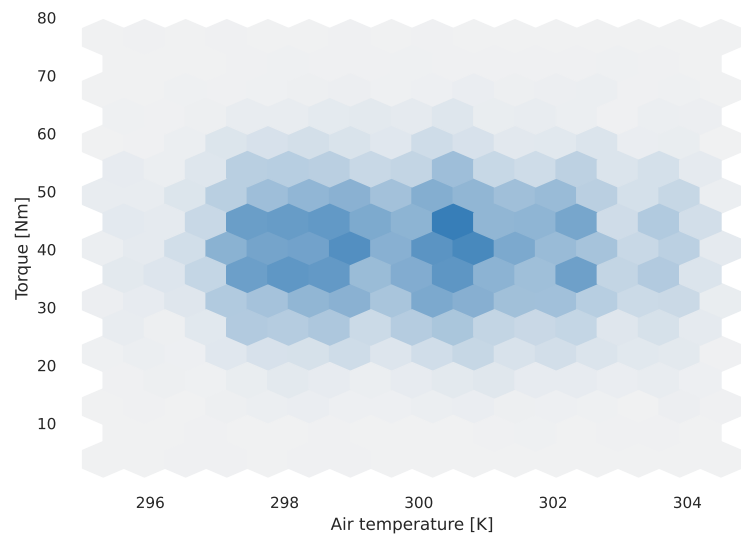
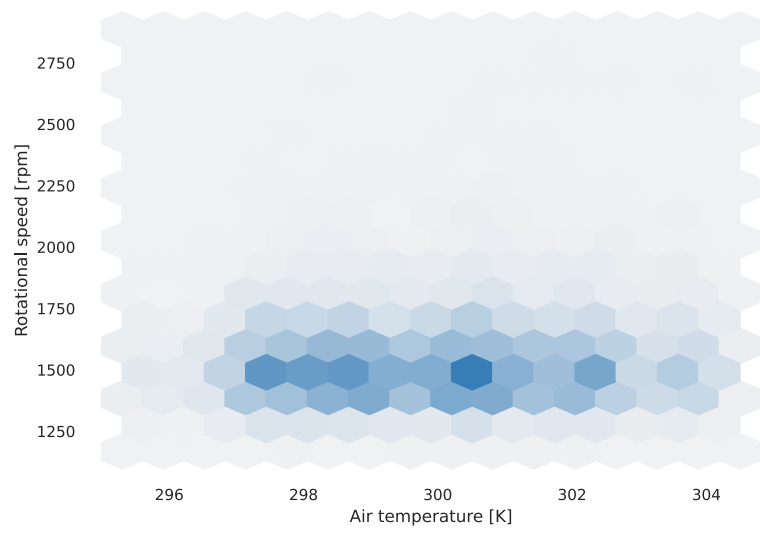
Most frequent character per block

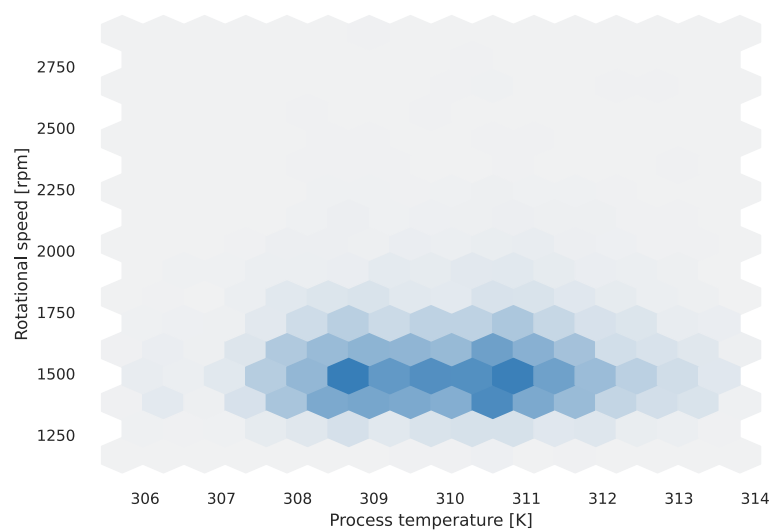
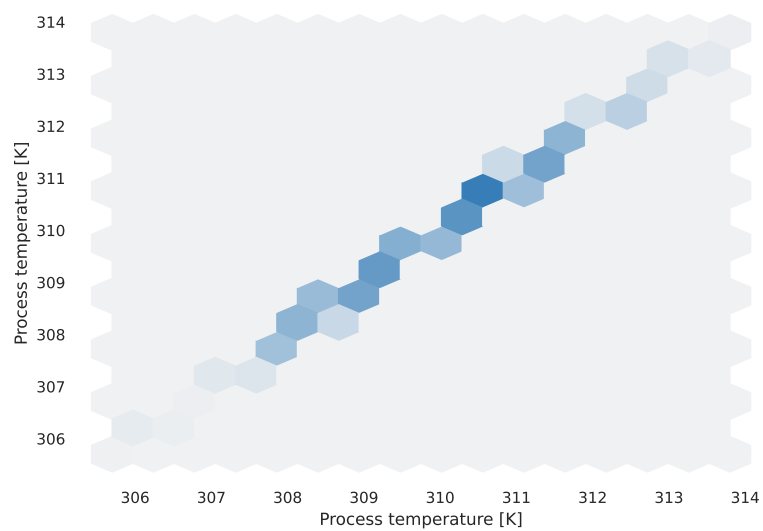
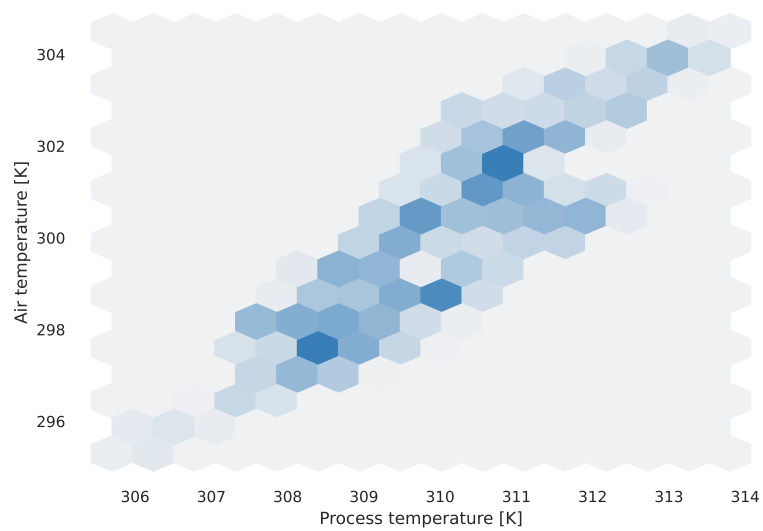
ASCII

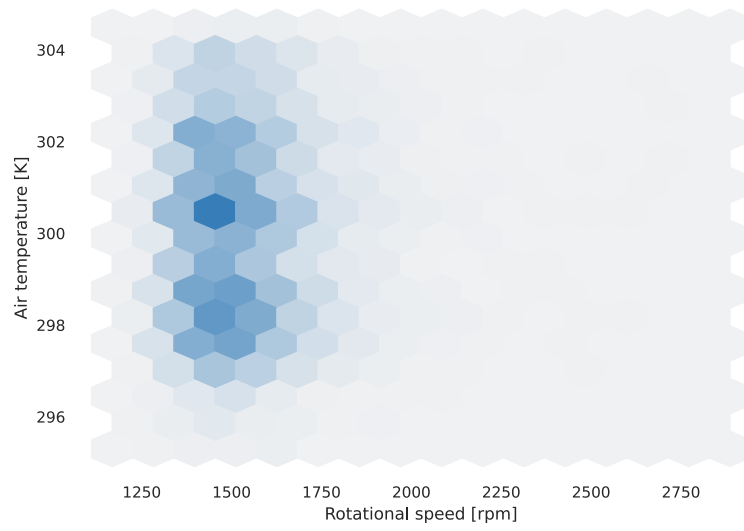
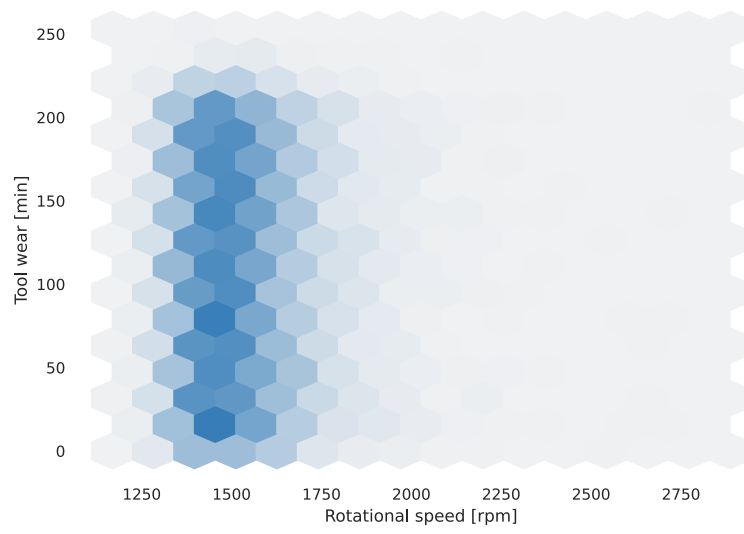
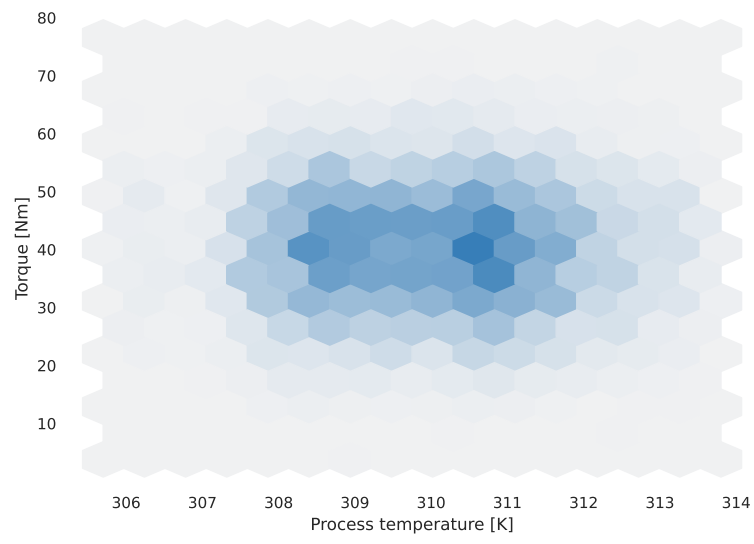
Value	Count	Frequency (%)
0	9661	96.6%
1	339	3.4%

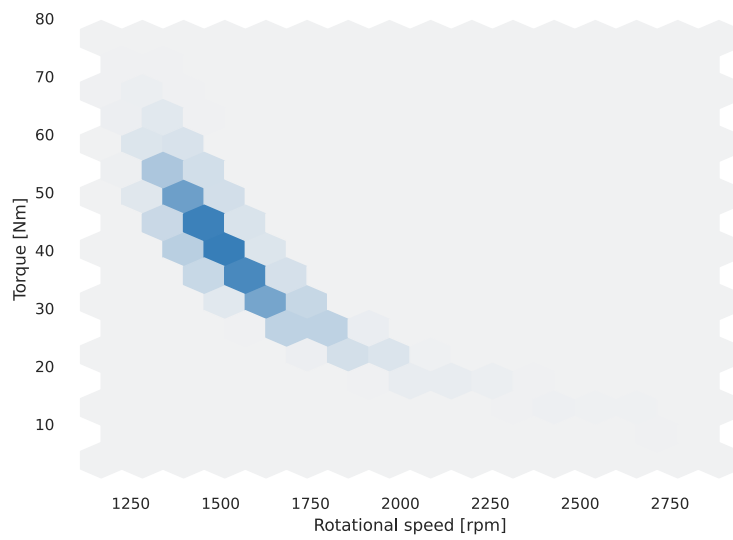
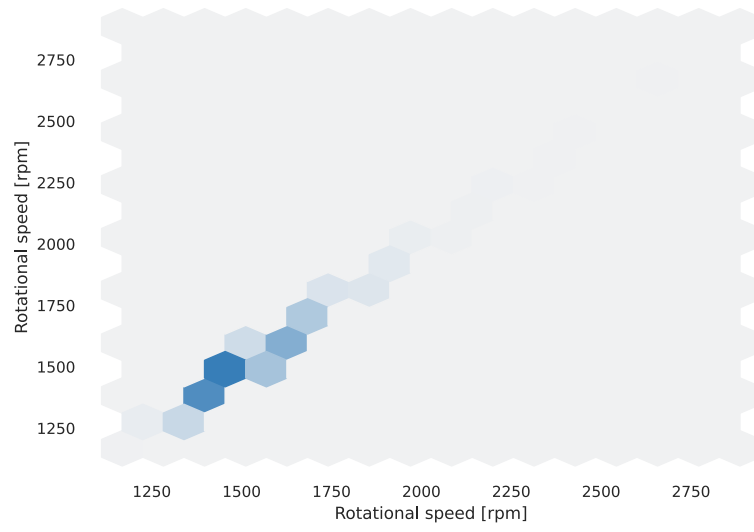
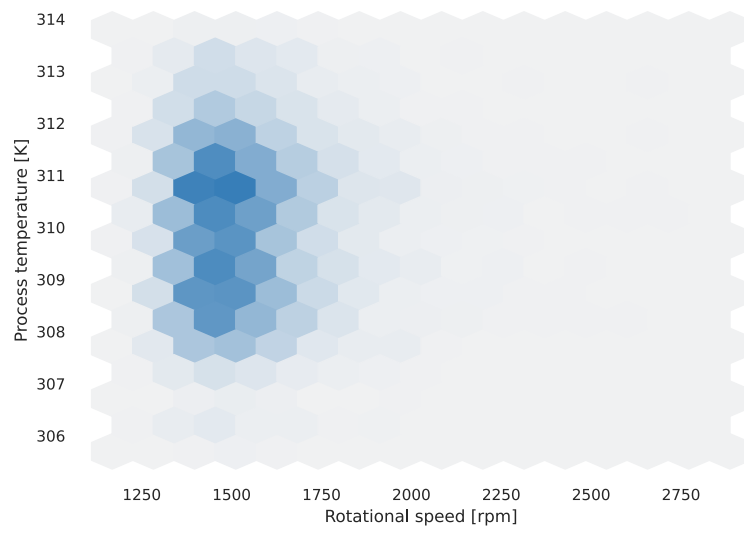
Interactions

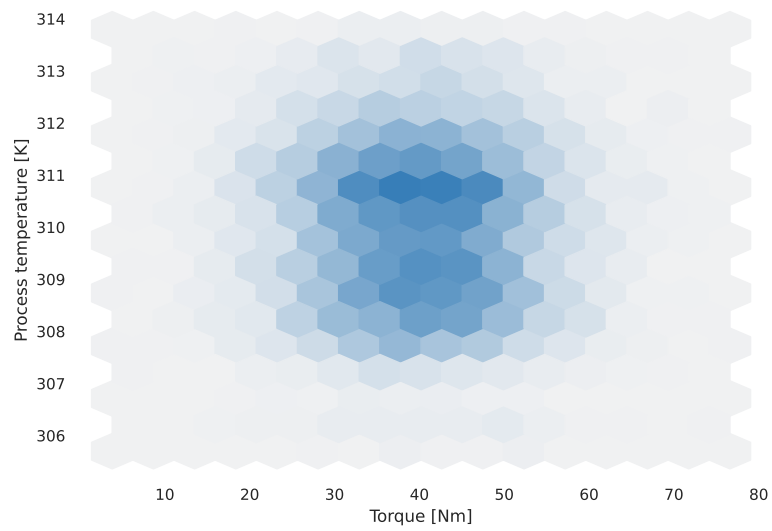
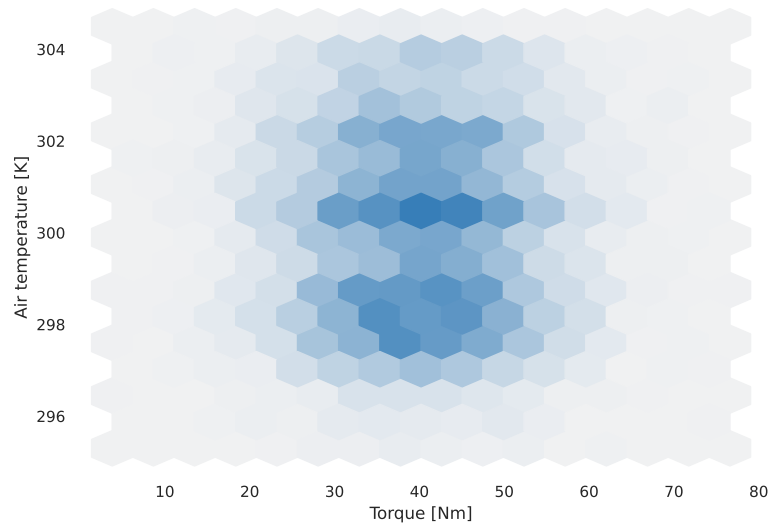
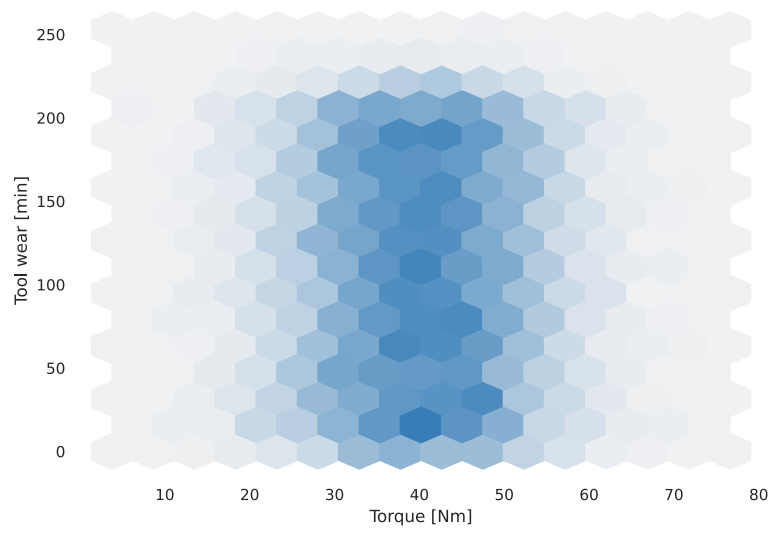


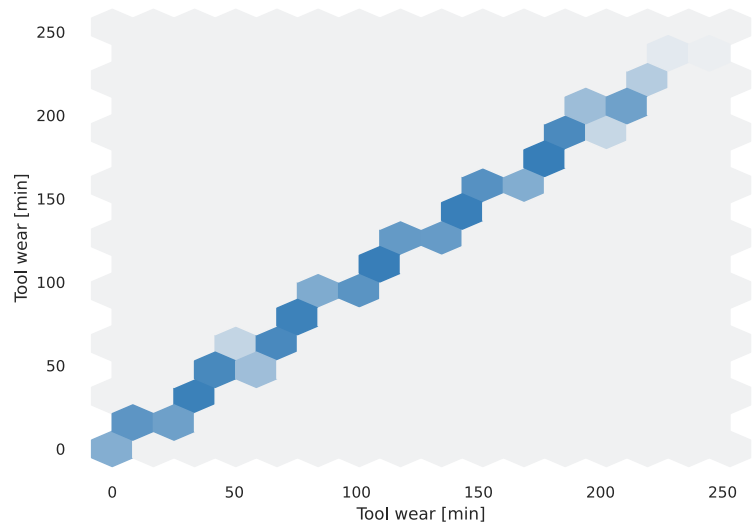
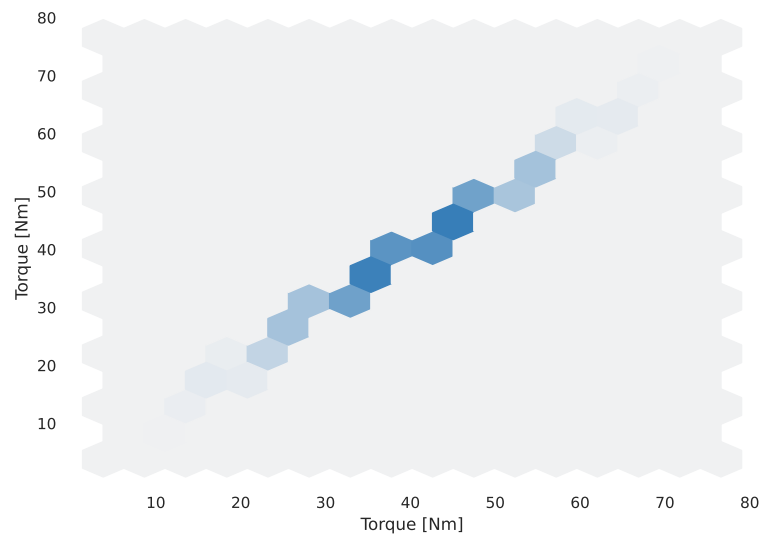
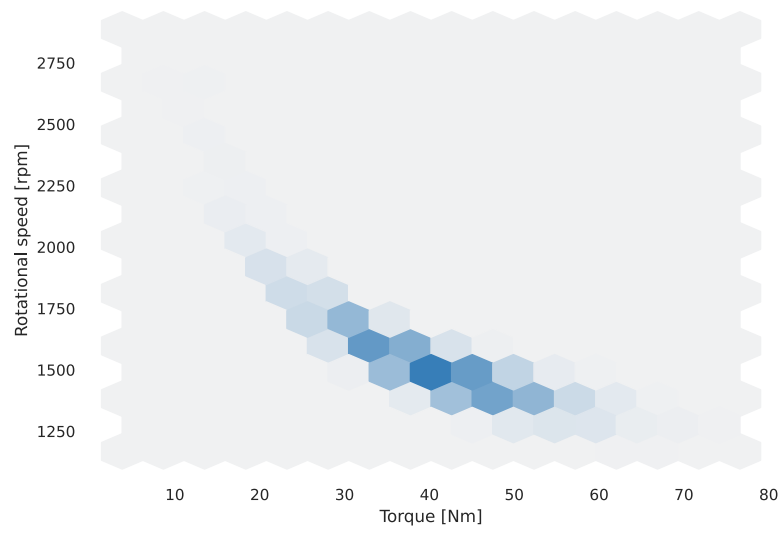


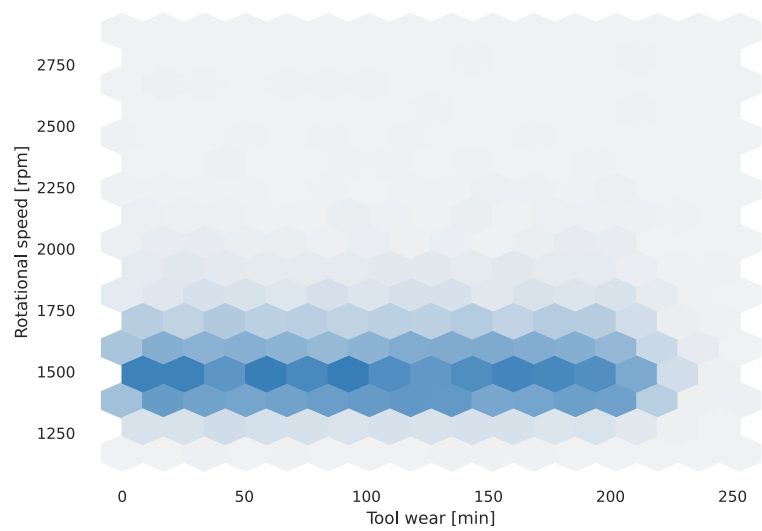
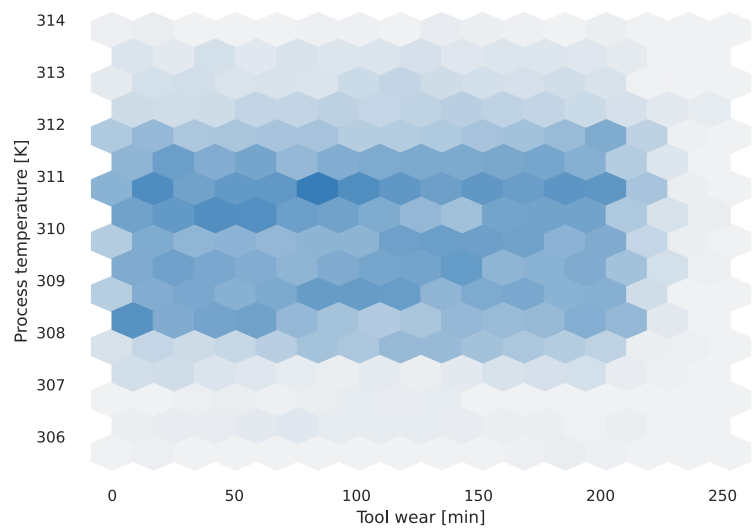
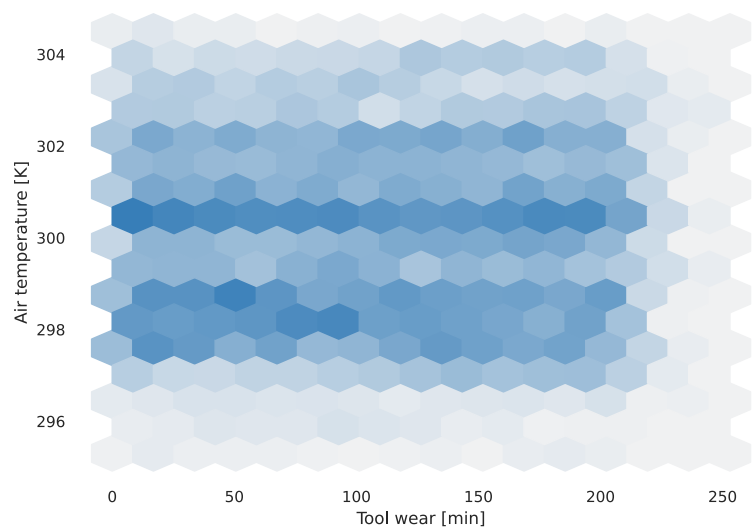


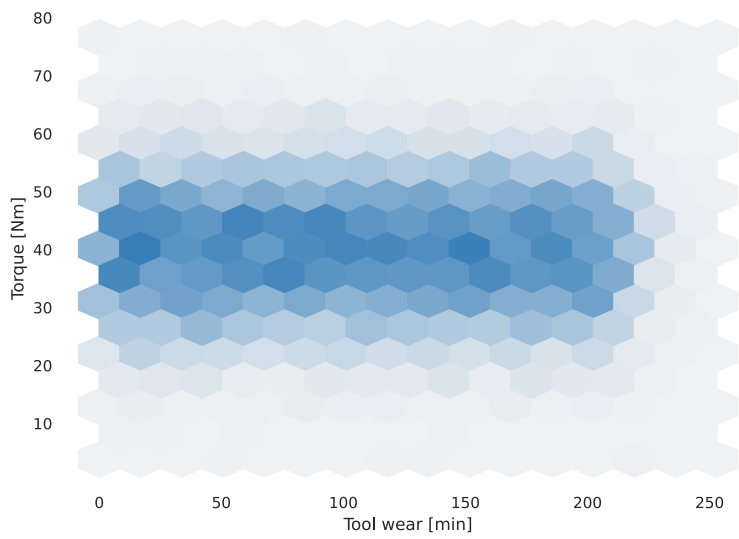




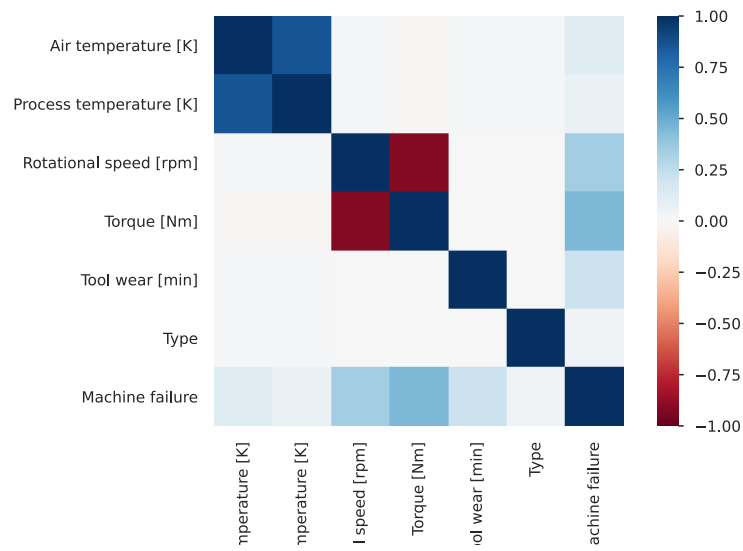








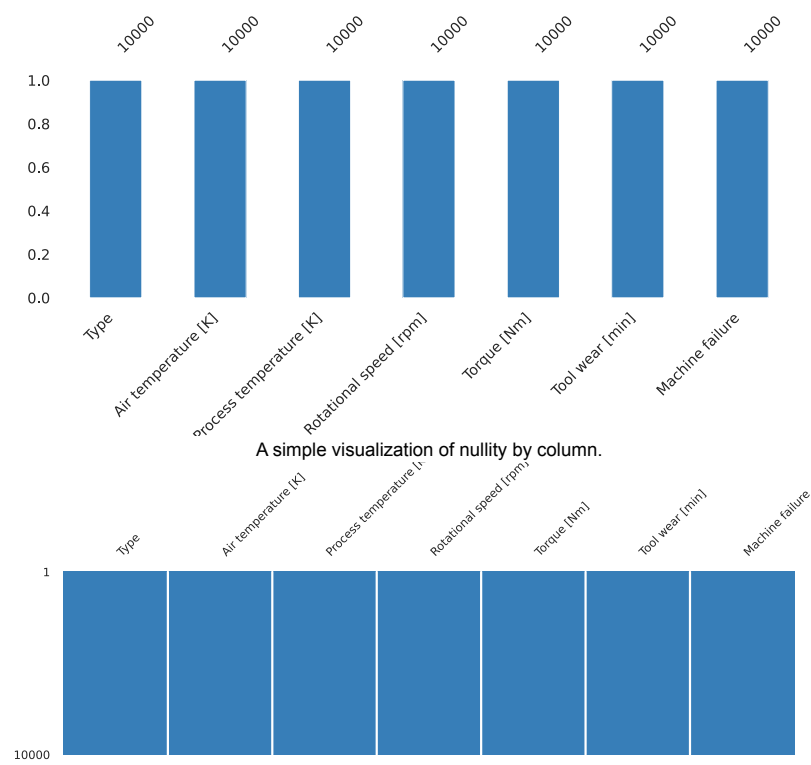
Correlations



	Air temperature [K]	Process temperature [K]	Rotational speed [rpm]	Torque [Nm]	Tool wear [min]	Type	Machine failure
Air temperature [K]	1.000	0.864	0.014	-0.012	0.013	0.014	0.123
Process temperature [K]	0.864	1.000	0.017	-0.014	0.014	0.014	0.067
Rotational speed [rpm]	0.014	0.017	1.000	-0.916	0.003	0.000	0.345
Torque [Nm]	-0.012	-0.014	-0.916	1.000	-0.003	-0.000	-0.345
Tool wear [min]	0.013	0.014	0.003	-0.003	1.000	0.000	0.000
Type	0.014	0.014	0.000	-0.000	0.000	1.000	0.000
Machine failure	0.123	0.067	0.345	-0.345	0.000	0.000	1.000



Missing values



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

Sample

Type		Air temperature [K]	Process temperature [K]	Rotational speed [rpm]	Torque [Nm]
0	M	298.10	308.60	1551	42.80
1	L	298.20	308.70	1408	46.30
2	L	298.10	308.50	1498	49.40
3	L	298.20	308.60	1433	39.50
4	L	298.20	308.70	1408	40.00
5	M	298.10	308.60	1425	41.90
6	L	298.10	308.60	1558	42.40
7	L	298.10	308.60	1527	40.20
8	M	298.30	308.70	1667	28.60
9	M	298.50	309.00	1741	28.00

Type		Air temperature [K]	Process temperature [K]	Rotational speed [rpm]	Torque [Nm]
9990	L	298.80	308.50	1527	36.20
9991	M	298.90	308.40	1827	26.10
9992	L	298.80	308.40	1484	39.20
9993	L	298.80	308.40	1401	47.30
9994	L	298.80	308.30	1634	27.90
9995	M	298.80	308.40	1604	29.50
9996	H	298.90	308.40	1632	31.80
9997	M	299.00	308.60	1645	33.40
9998	H	299.00	308.70	1408	48.50

	Type	Air temperature [K]	Process temperature [K]	Rotational speed [rpm]	Torque [Nm]
9999	M	299.00	308.70	1500	40.20

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

