Overview

Dataset statistics

Number of variables	8
Number of observations	900
Missing cells	0
Missing cells (%)	0.0%
Duplicate rows	0
Duplicate rows (%)	0.0%
Total size in memory	56.4 KiB
Average record size in memory	64.1 B

Variable types

Numeric	7
Categorical	1

Alerts

`Area` is highly overall correlated with `MajorAxisLength` and 4 other fields (MajorAxisLength, MinorAxisLength, ConvexArea, Perimeter, Class)	High correlation
`MajorAxisLength` is highly overall correlated with `Area` and 5 other fields (Area, MinorAxisLength, Eccentricity, ConvexArea, Perimeter, Class)	High correlation
`MinorAxisLength` is highly overall correlated with `Area` and 3 other fields (Area, MajorAxisLength, ConvexArea, Perimeter)	High correlation
`Eccentricity` is highly overall correlated with `MajorAxisLength` and 1 other fields (MajorAxisLength, Perimeter)	High correlation
`ConvexArea` is highly overall correlated with `Area` and 4 other fields (Area, MajorAxisLength, MinorAxisLength, Perimeter, Class)	High correlation
`Perimeter` is highly overall correlated with `Area` and 5 other fields (Area, MajorAxisLength, MinorAxisLength, Eccentricity, ConvexArea, Class)	High correlation
`Class` is highly overall correlated with `Area` and 3 other fields (Area, MajorAxisLength, ConvexArea, Perimeter)	High correlation
`Class` is uniformly distributed	Uniform
`Area` has unique values	Unique
`MajorAxisLength` has unique values	Unique
`MinorAxisLength` has unique values	Unique
`Eccentricity` has unique values	Unique
`Extent` has unique values	Unique
`Perimeter` has unique values	Unique

Reproduction

Analysis started	2023-07-04 21:33:29.477110
Analysis finished	2023-07-04 21:33:41.209134
Duration	11.73 seconds
Software version	pandas-profiling vv3.6.2 (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%22%22%2C%20%22creator%22%3A%20%22%22%2C%20%22author%22%3A%20%22%22%2C%20%22copyright_holder%22%3A%20%22%22%2C%20%22copyright_year%22%3A%20%22%22%2C%20%22url%22%3A%20%22%22%7D%2C%20%22variables%22%3A%20%7B%22descriptions%22%3A%20%7B%7D%7D%2C%20%22infer_dtypes%22%3A%20true%2C%20%22show_variable_description%22%3A%20true%2C%20%22pool_size%22%3A%200%2C%20%22progress_bar%22%3A%20true%2C%20%22vars%22%3A%20%7B%22num%22%3A%20%7B%22quantiles%22%3A%20%5B0.05%2C%200.25%2C%200.5%2C%200.75%2C%200.95%5D%2C%20%22skewness_threshold%22%3A%2020%2C%20%22low_categorical_threshold%22%3A%205%2C%20%22chi_squared_threshold%22%3A%200.999%7D%2C%20%22cat%22%3A%20%7B%22length%22%3A%20true%2C%20%22characters%22%3A%20true%2C%20%22words%22%3A%20true%2C%20%22cardinality_threshold%22%3A%2050%2C%20%22imbalance_threshold%22%3A%200.5%2C%20%22n_obs%22%3A%205%2C%20%22chi_squared_threshold%22%3A%200.999%2C%20%22coerce_str_to_date%22%3A%20false%2C%20%22redact%22%3A%20false%2C%20%22histogram_largest%22%3A%2050%2C%20%22stop_words%22%3A%20%5B%5D%7D%2C%20%22image%22%3A%20%7B%22active%22%3A%20false%2C%20%22exif%22%3A%20true%2C%20%22hash%22%3A%20true%7D%2C%20%22bool%22%3A%20%7B%22n_obs%22%3A%203%2C%20%22imbalance_threshold%22%3A%200.5%2C%20%22mappings%22%3A%20%7B%22t%22%3A%20true%2C%20%22f%22%3A%20false%2C%20%22yes%22%3A%20true%2C%20%22no%22%3A%20false%2C%20%22y%22%3A%20true%2C%20%22n%22%3A%20false%2C%20%22true%22%3A%20true%2C%20%22false%22%3A%20false%7D%7D%2C%20%22path%22%3A%20%7B%22active%22%3A%20false%7D%2C%20%22file%22%3A%20%7B%22active%22%3A%20false%7D%2C%20%22url%22%3A%20%7B%22active%22%3A%20false%7D%2C%20%22timeseries%22%3A%20%7B%22active%22%3A%20false%2C%20%22sortby%22%3A%20null%2C%20%22autocorrelation%22%3A%200.7%2C%20%22lags%22%3A%20%5B1%2C%207%2C%2012%2C%2024%2C%2030%5D%2C%20%22significance%22%3A%200.05%2C%20%22pacf_acf_lag%22%3A%20100%7D%7D%2C%20%22sort%22%3A%20null%2C%20%22missing_diagrams%22%3A%20%7B%22bar%22%3A%20true%2C%20%22matrix%22%3A%20true%2C%20%22heatmap%22%3A%20true%7D%2C%20%22correlations%22%3A%20%7B%22auto%22%3A%20%7B%22key%22%3A%20%22auto%22%2C%20%22calculate%22%3A%20true%2C%20%22warn_high_correlations%22%3A%2010%2C%20%22threshold%22%3A%200.5%2C%20%22n_bins%22%3A%2010%7D%7D%2C%20%22correlation_table%22%3A%20true%2C%20%22interactions%22%3A%20%7B%22continuous%22%3A%20true%2C%20%22targets%22%3A%20%5B%5D%7D%2C%20%22categorical_maximum_correlation_distinct%22%3A%20100%2C%20%22memory_deep%22%3A%20false%2C%20%22plot%22%3A%20%7B%22missing%22%3A%20%7B%22force_labels%22%3A%20true%2C%20

%22cmap%22%3A%20%22RdBu%22%7D%2C%20%22image_format%22%3A
%20%22svg%22%2C%20%22correlation%22%3A%20%7B%22cmap%22%3A
%20%22RdBu%22%2C%20%22bad%22%3A%20%22%23000000%22%7D
%2C%20%22dpi%22%3A%20800%2C%20%22histogram%22%3A%20%7B
%22bins%22%3A%2050%2C%20%22max_bins%22%3A%20250%2C%20
%22x_axis_labels%22%3A%20true%7D%2C%20%22scatter_threshold%22%3A
%201000%2C%20%22cat_freq%22%3A%20%7B%22show%22%3A%20true
%2C%20%22type%22%3A%20%22bar%22%2C%20%22max_unique%22%3A
%2010%2C%20%22colors%22%3A%20null%7D%7D%2C%20%22duplicates
%22%3A%20%7B%22head%22%3A%2010%2C%20%22key%22%3A%20%22
%23%20duplicates%22%7D%2C%20%22samples%22%3A%20%7B%22head
%22%3A%2010%2C%20%22tail%22%3A%2010%2C%20%22random%22%3A
%200%7D%2C%20%22reject_variables%22%3A%20true%2C%20
%22n_obs_unique%22%3A%2010%2C%20%22n_freq_table_max%22%3A
%2010%2C%20%22n_extreme_obs%22%3A%2010%2C%20%22report%22%3A
%20%7B%22precision%22%3A%208%7D%2C%20%22html%22%3A%20%7B
%22style%22%3A%20%7B%22primary_colors%22%3A%20%5B%22%23377eb8
%22%2C%20%22%23e41a1c%22%2C%20%22%234daf4a%22%5D%2C%20
%22logo%22%3A%20%22%22%2C%20%22theme%22%3A%20null%7D%2C
%20%22navbar_show%22%3A%20true%2C%20%22minify_html%22%3A%20true
%2C%20%22use_local_assets%22%3A%20true%2C%20%22inline%22%3A
%20true%2C%20%22assets_prefix%22%3A%20null%2C%20%22assets_path
%22%3A%20null%2C%20%22full_width%22%3A%20false%7D%2C
%20%22notebook%22%3A%20%7B%22iframe%22%3A%20%7B%22height
%22%3A%20%22800px%22%2C%20%22width%22%3A%20%22100%25%22
%2C%20%22attribute%22%3A%20%22srcdoc%22%7D%7D%7D)

Variables

Select Columns

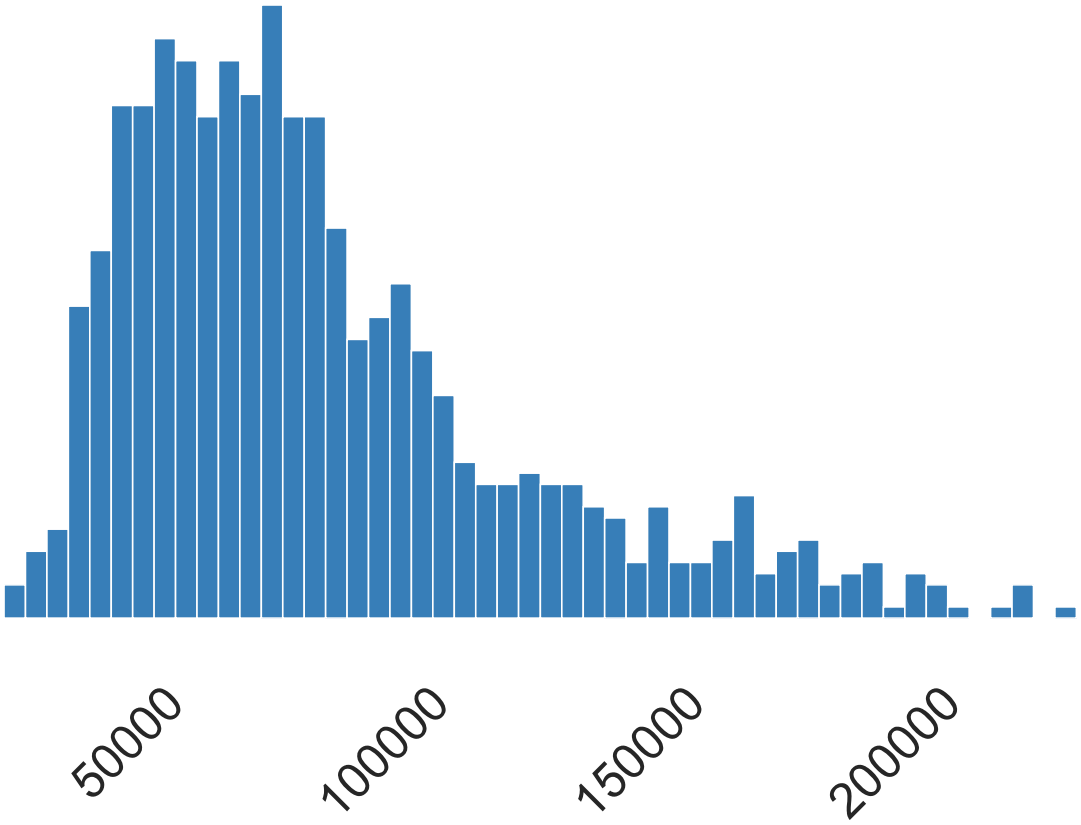
Area

Real number (ℝ)

HIGH CORRELATION (This variable has a high overall correlation with 5 fields: MajorAxisLength, MinorAxisLength, ConvexArea, Perimeter, Class) UNIQUE

Distinct	900
Distinct (%)	100.0%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Mean	87804.128

Minimum	25387
Maximum	235047
Zeros	0
Zeros (%)	0.0%
Negative	0
Negative (%)	0.0%
Memory size	7.2 KiB



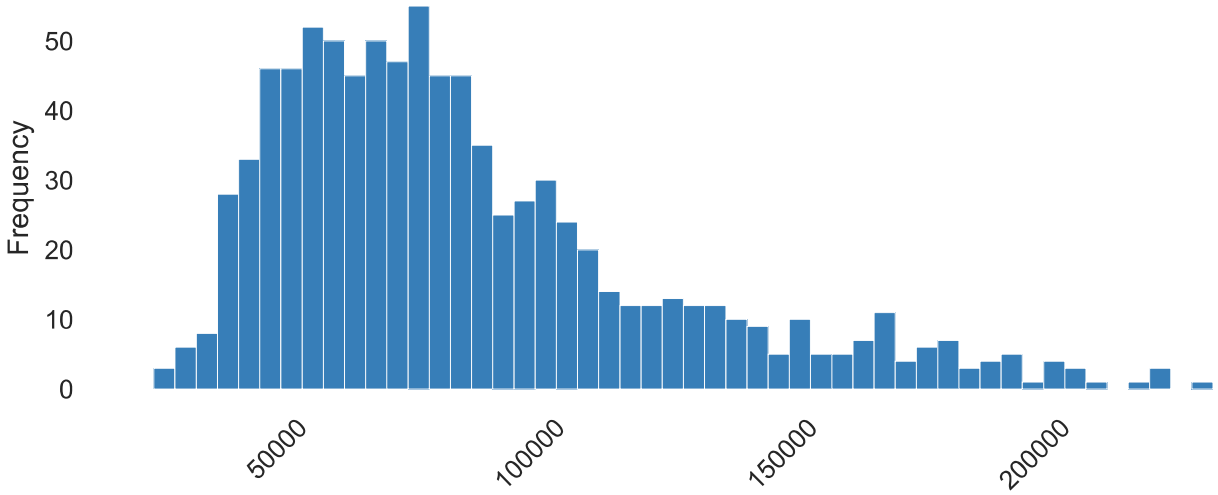
Quantile statistics

Minimum	25387
5-th percentile	42242.95
Q1	59348
median	78902
Q3	105028.25
95-th percentile	171256.4
Maximum	235047
Range	209660
Interquartile range (IQR)	45680.25

Descriptive statistics

Standard deviation	39002.111

Coefficient of variation (CV)	0.44419451
Kurtosis	1.0740729
Mean	87804.128
Median Absolute Deviation (MAD)	21615
Skewness	1.1752374
Sum	79023715
Variance	1.5211647×10^9
Monotonicity	Not monotonic



Histogram with fixed size bins (bins=50)

Value	Count	Frequency (%)
87524	1	0.1%
118043	1	0.1%
124166	1	0.1%
133784	1	0.1%
80481	1	0.1%
63491	1	0.1%
109791	1	0.1%
89236	1	0.1%
152267	1	0.1%
89721	1	0.1%
Other values (890)	890	98.9%

Value	Count	Frequency (%)
25387	1	0.1%
26908	1	0.1%
28216	1	0.1%
31237	1	0.1%
31275	1	0.1%
32097	1	0.1%
33565	1	0.1%
33615	1	0.1%
33662	1	0.1%
34559	1	0.1%

Value	Count	Frequency (%)
235047	1	0.1%
225043	1	0.1%
223075	1	0.1%
222915	1	0.1%
218459	1	0.1%
210923	1	0.1%
208264	1	0.1%
206720	1	0.1%
206689	1	0.1%
205497	1	0.1%

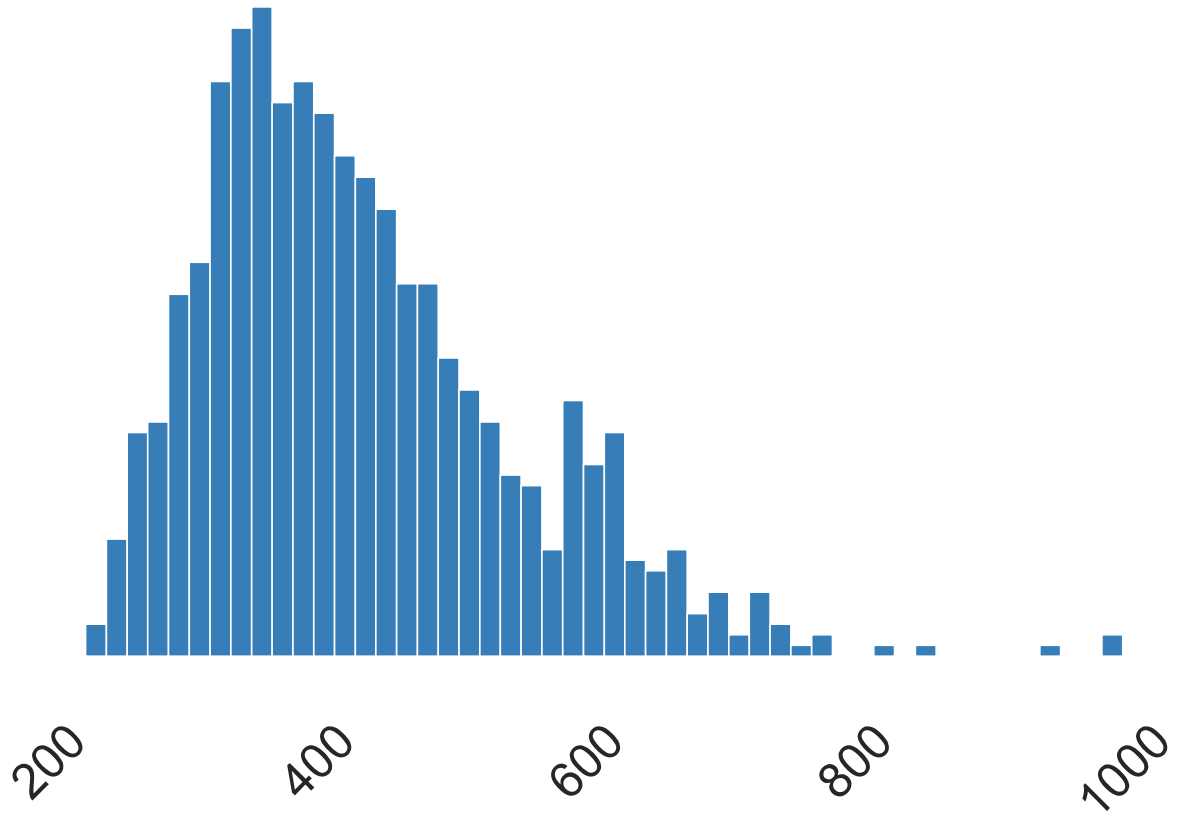
MajorAxisLength

Real number (ℝ)

HIGH CORRELATION (This variable has a high overall correlation with 6 fields: Area, MinorAxisLength, Eccentricity, ConvexArea, Perimeter, Class) UNIQUE

Distinct	900
Distinct (%)	100.0%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Mean	430.92995

Minimum	225.62954
Maximum	997.29194
Zeros	0
Zeros (%)	0.0%
Negative	0
Negative (%)	0.0%
Memory size	7.2 KiB



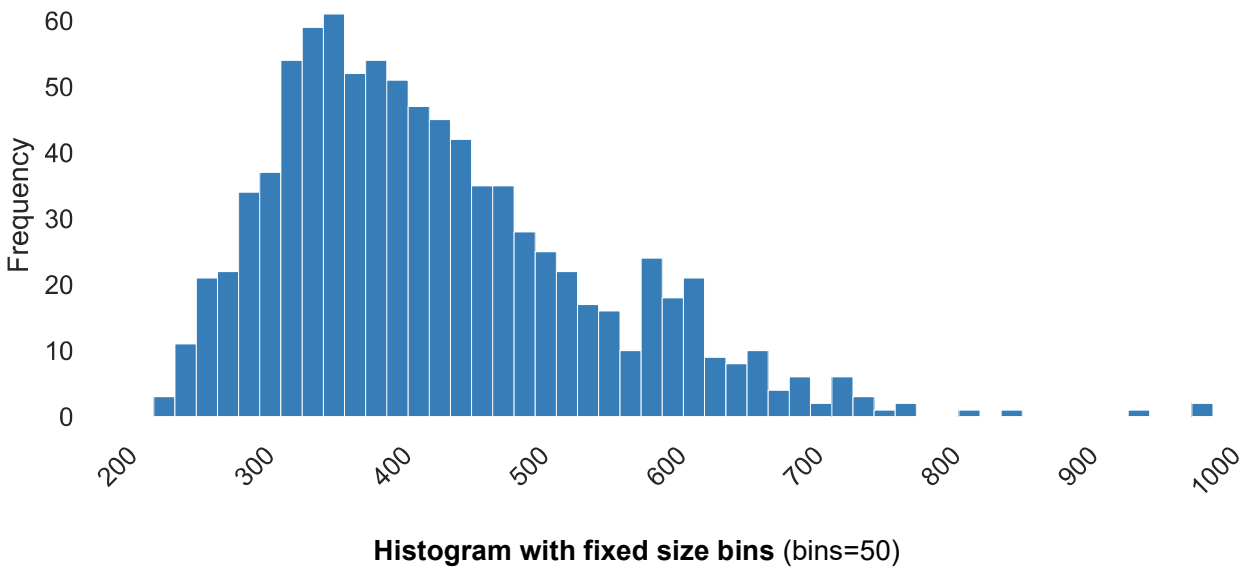
Quantile statistics

Minimum	225.62954
5-th percentile	280.30492
Q1	345.4429
median	407.80395
Q3	494.18701
95-th percentile	648.23929
Maximum	997.29194
Range	771.6624
Interquartile range (IQR)	148.74412

Descriptive statistics

Standard deviation	116.03512

Coefficient of variation (CV)	0.26926678
Kurtosis	1.326808
Mean	430.92995
Median Absolute Deviation (MAD)	71.099522
Skewness	0.98954413
Sum	387836.96
Variance	13464.149
Monotonicity	Not monotonic



Histogram with fixed size bins (bins=50)

Value	Count	Frequency (%)
442.2460114	1	0.1%
493.7652445	1	0.1%
525.9453568	1	0.1%
581.2869477	1	0.1%
481.0639534	1	0.1%
326.6328824	1	0.1%
477.0335016	1	0.1%
389.6816918	1	0.1%
598.9676358	1	0.1%
530.1565743	1	0.1%
Other values (890)	890	98.9%

Value	Count	Frequency (%)
225.629541	1	0.1%
227.2937917	1	0.1%
232.4278475	1	0.1%
243.0382802	1	0.1%
245.4012948	1	0.1%
245.7557814	1	0.1%
246.7636115	1	0.1%
249.7402266	1	0.1%
251.1333844	1	0.1%
251.7422462	1	0.1%

Value	Count	Frequency (%)
997.2919406	1	0.1%
984.0454912	1	0.1%
949.6626718	1	0.1%
843.9566534	1	0.1%
820.724022	1	0.1%
772.956877	1	0.1%
769.4251488	1	0.1%
755.0129141	1	0.1%
746.145341	1	0.1%
740.1087099	1	0.1%

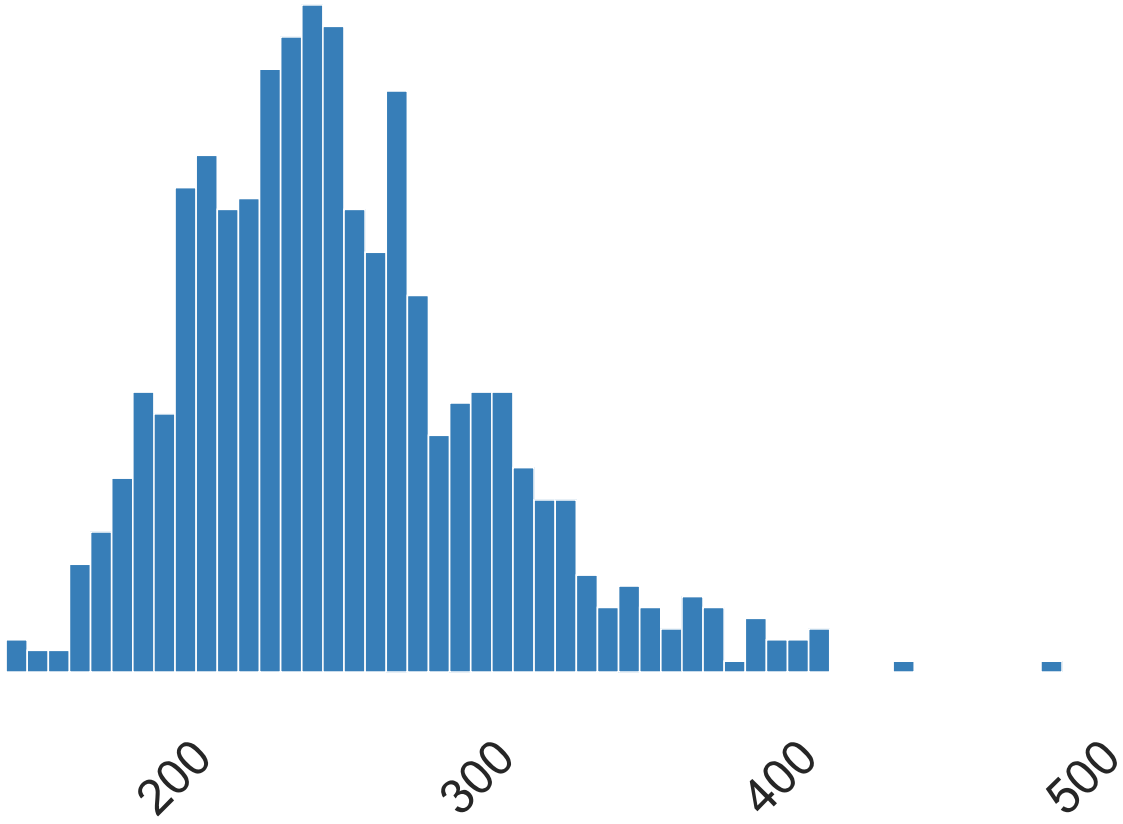
MinorAxisLength

Real number (ℝ)

HIGH CORRELATION (This variable has a high overall correlation with 4 fields: Area, MajorAxisLength, ConvexArea, Perimeter) UNIQUE

Distinct	900
Distinct (%)	100.0%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Mean	254.48813

Minimum	143.71087
Maximum	492.27528
Zeros	0
Zeros (%)	0.0%
Negative	0
Negative (%)	0.0%
Memory size	7.2 KiB



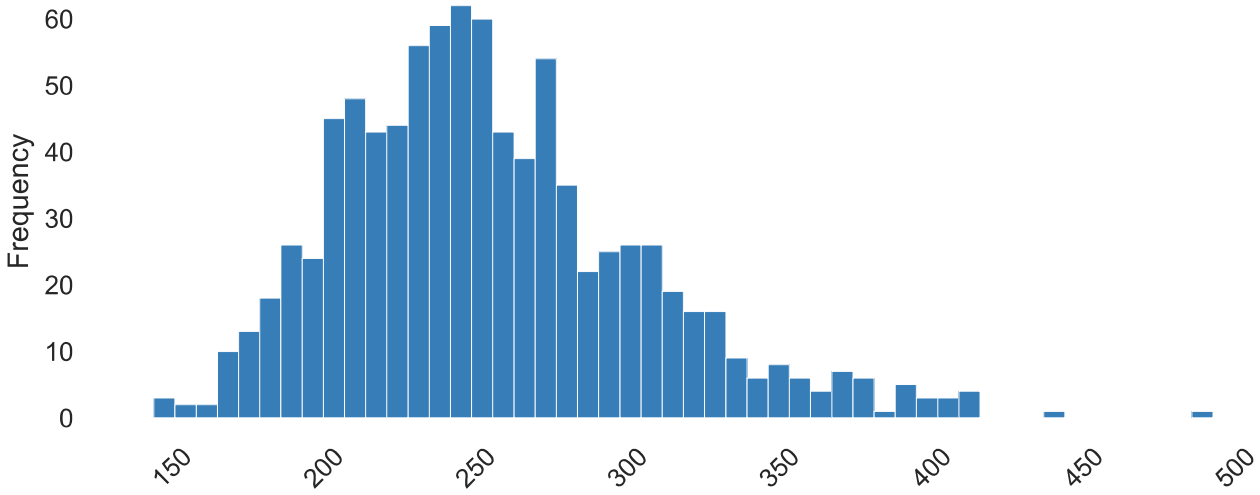
Quantile statistics

Minimum	143.71087
5-th percentile	184.7728
Q1	219.11113
median	247.84841
Q3	279.88857
95-th percentile	348.351
Maximum	492.27528
Range	348.56441
Interquartile range (IQR)	60.777448

Descriptive statistics

Standard deviation	49.988902

Coefficient of variation (CV)	0.19642921
Kurtosis	0.95391521
Mean	254.48813
Median Absolute Deviation (MAD)	29.982991
Skewness	0.80004936
Sum	229039.32
Variance	2498.8903
Monotonicity	Not monotonic



Histogram with fixed size bins (bins=50)

Value	Count	Frequency (%)
253.291155	1	0.1%
308.0001269	1	0.1%
304.1564686	1	0.1%
300.635686	1	0.1%
217.561151	1	0.1%
248.3224493	1	0.1%
294.5338471	1	0.1%
295.3319356	1	0.1%
331.2490172	1	0.1%
223.4999334	1	0.1%
Other values (890)	890	98.9%

Value	Count	Frequency (%)
143.7108718	1	0.1%
144.618672	1	0.1%
150.2455819	1	0.1%
155.0257428	1	0.1%
156.3955455	1	0.1%
157.9904179	1	0.1%
161.7023657	1	0.1%
166.5935504	1	0.1%
167.3283344	1	0.1%
167.7084908	1	0.1%

Value	Count	Frequency (%)
492.2752785	1	0.1%
440.4971275	1	0.1%
414.1883255	1	0.1%
413.9274732	1	0.1%
412.3828168	1	0.1%
411.810369	1	0.1%
408.5356188	1	0.1%
403.7193275	1	0.1%
402.2832708	1	0.1%
400.8016603	1	0.1%

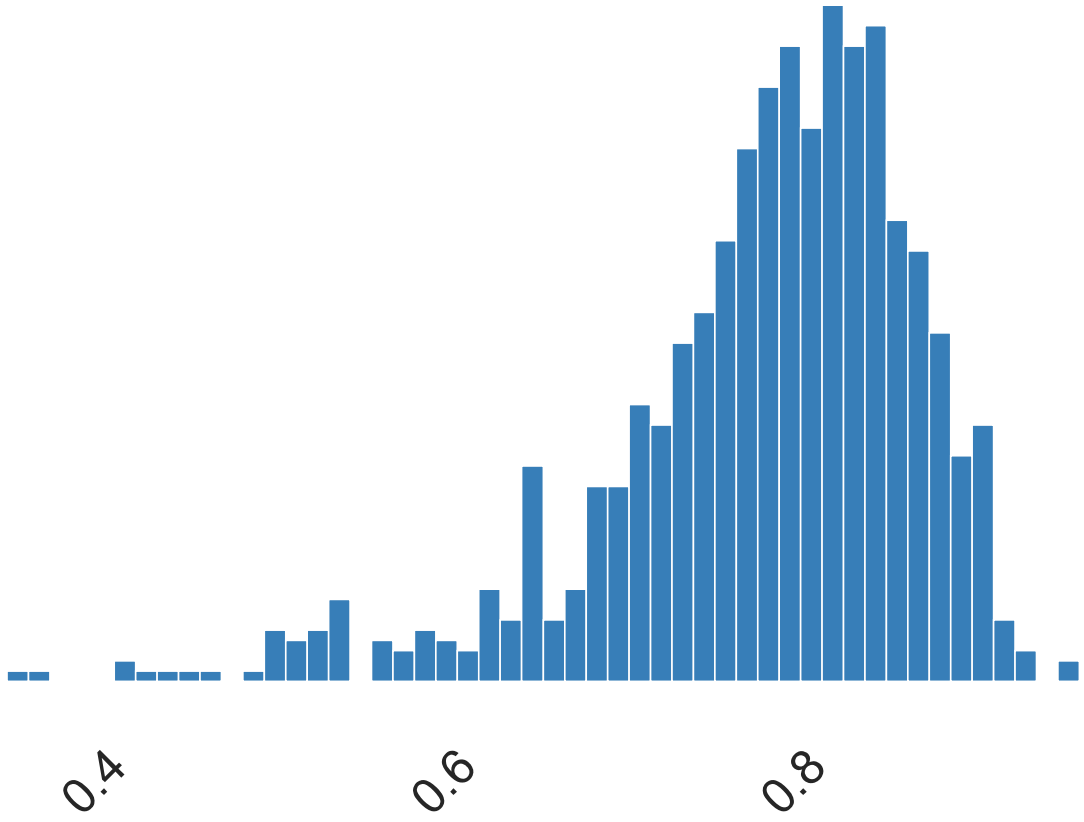
Eccentricity

Real number (ℝ)

HIGH CORRELATION (This variable has a high overall correlation with 2 fields: MajorAxisLength, Perimeter) UNIQUE

Distinct	900
Distinct (%)	100.0%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Mean	0.78154215

Minimum	0.34872964
Maximum	0.96212444
Zeros	0
Zeros (%)	0.0%
Negative	0
Negative (%)	0.0%
Memory size	7.2 KiB



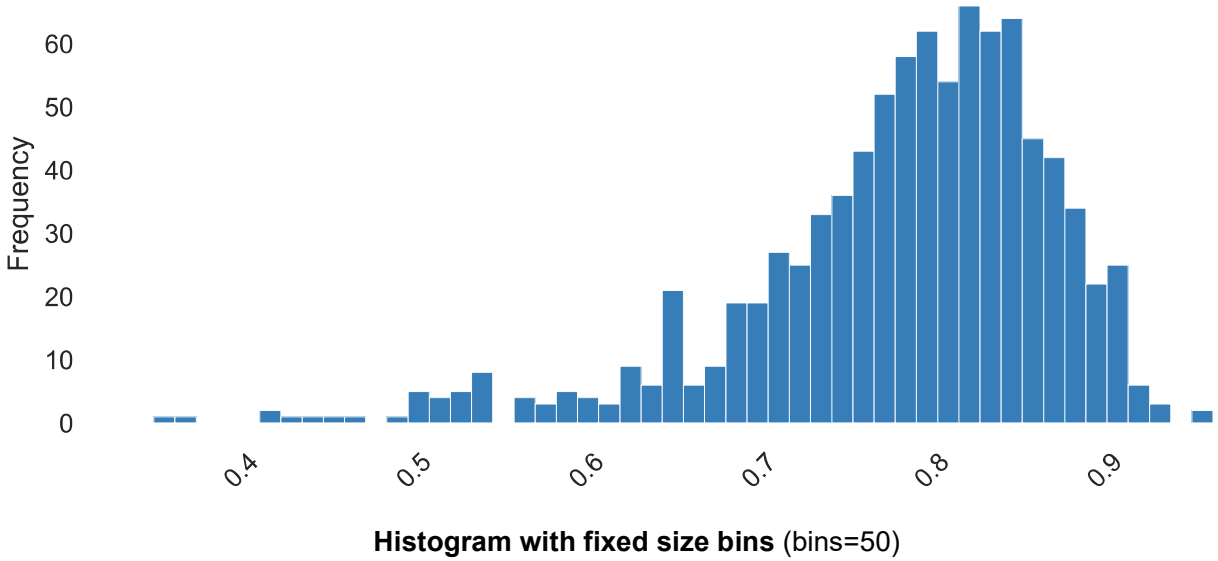
Quantile statistics

Minimum	0.34872964
5-th percentile	0.60293774
Q1	0.74176625
median	0.79884604
Q3	0.84257102
95-th percentile	0.89324309
Maximum	0.96212444
Range	0.6133948
Interquartile range (IQR)	0.10080477

Descriptive statistics

Standard deviation	0.09031841

Coefficient of variation (CV)	0.11556435
Kurtosis	2.4921214
Mean	0.78154215
Median Absolute Deviation (MAD)	0.049858966
Skewness	-1.3275032
Sum	703.38794
Variance	0.0081574152
Monotonicity	Not monotonic



Histogram with fixed size bins (bins=50)

Value	Count	Frequency (%)
0.819738392	1	0.1%
0.781601193	1	0.1%
0.815821133	1	0.1%
0.855870727	1	0.1%
0.891891237	1	0.1%
0.649631331	1	0.1%
0.786627389	1	0.1%
0.652394055	1	0.1%
0.833159184	1	0.1%
0.90679426	1	0.1%
Other values (890)	890	98.9%

Value	Count	Frequency (%)
0.348729642	1	0.1%
0.369212459	1	0.1%
0.418381972	1	0.1%
0.419753707	1	0.1%
0.432307345	1	0.1%
0.44495009	1	0.1%
0.458544698	1	0.1%
0.460121209	1	0.1%
0.492086934	1	0.1%
0.496118267	1	0.1%

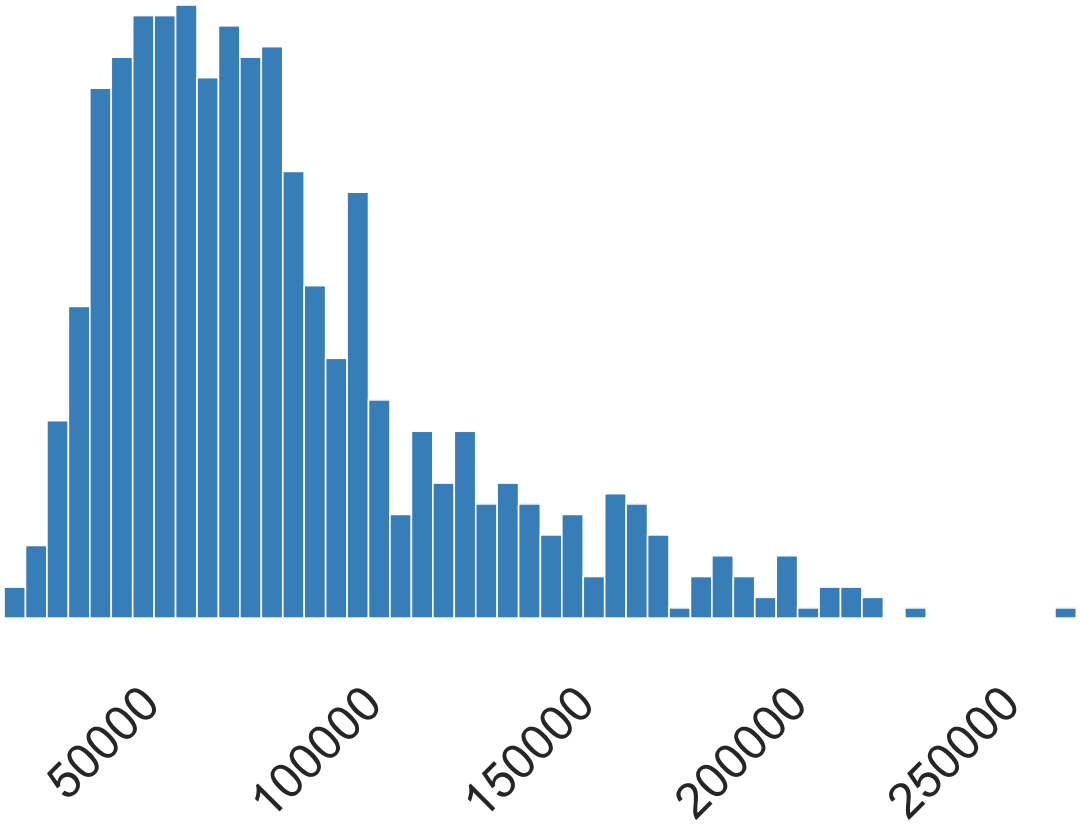
Value	Count	Frequency (%)
0.96212444	1	0.1%
0.951082244	1	0.1%
0.928093615	1	0.1%
0.927737116	1	0.1%
0.925634934	1	0.1%
0.923770364	1	0.1%
0.922383026	1	0.1%
0.921939284	1	0.1%
0.917708543	1	0.1%
0.914840211	1	0.1%

ConvexArea

Real number (ℝ)

Distinct	896
Distinct (%)	99.6%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Mean	91186.09

Minimum	26139
Maximum	278217
Zeros	0
Zeros (%)	0.0%
Negative	0
Negative (%)	0.0%
Memory size	7.2 KiB



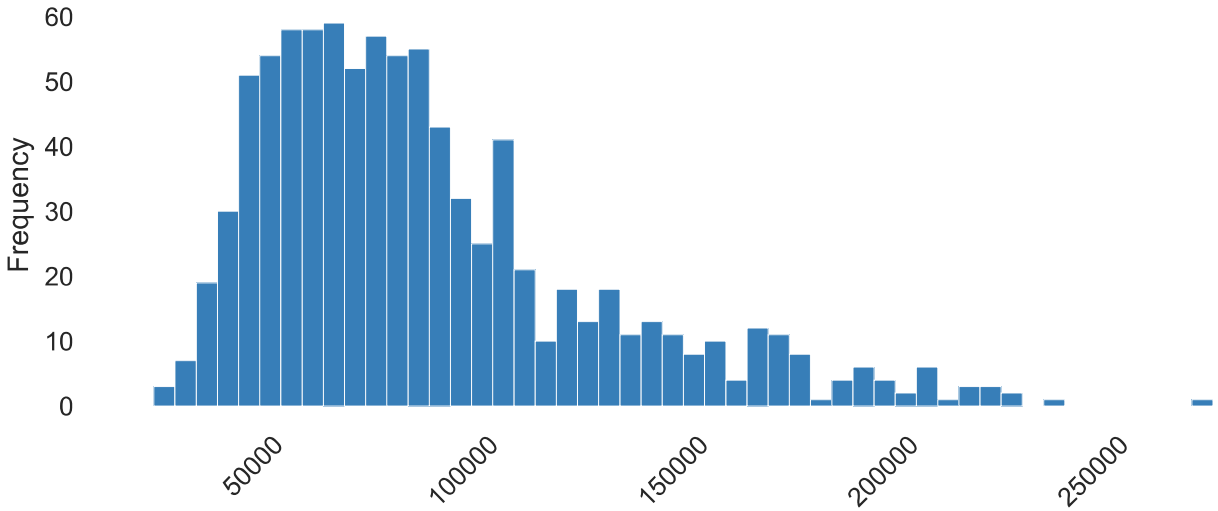
Quantile statistics

Minimum	26139
5-th percentile	44044.6
Q1	61513.25
median	81651
Q3	108375.75
95-th percentile	174766.05
Maximum	278217
Range	252078
Interquartile range (IQR)	46862.5

Descriptive statistics

Standard deviation	40769.29

Coefficient of variation (CV)	0.44709988
Kurtosis	1.427258
Mean	91186.09
Median Absolute Deviation (MAD)	22667.5
Skewness	1.242904
Sum	82067481
Variance	1.662135×10^9
Monotonicity	Not monotonic



Histogram with fixed size bins (bins=50)

Value	Count	Frequency (%)
70719	2	0.2%
81718	2	0.2%
92317	2	0.2%
49996	2	0.2%
58632	1	0.1%
64892	1	0.1%
112766	1	0.1%
92014	1	0.1%
157353	1	0.1%
95252	1	0.1%
Other values (886)	886	98.4%

Value	Count	Frequency (%)
26139	1	0.1%
28607	1	0.1%
30316	1	0.1%
32564	1	0.1%
33540	1	0.1%
33699	1	0.1%
34787	1	0.1%
35376	1	0.1%
35794	1	0.1%
35824	1	0.1%

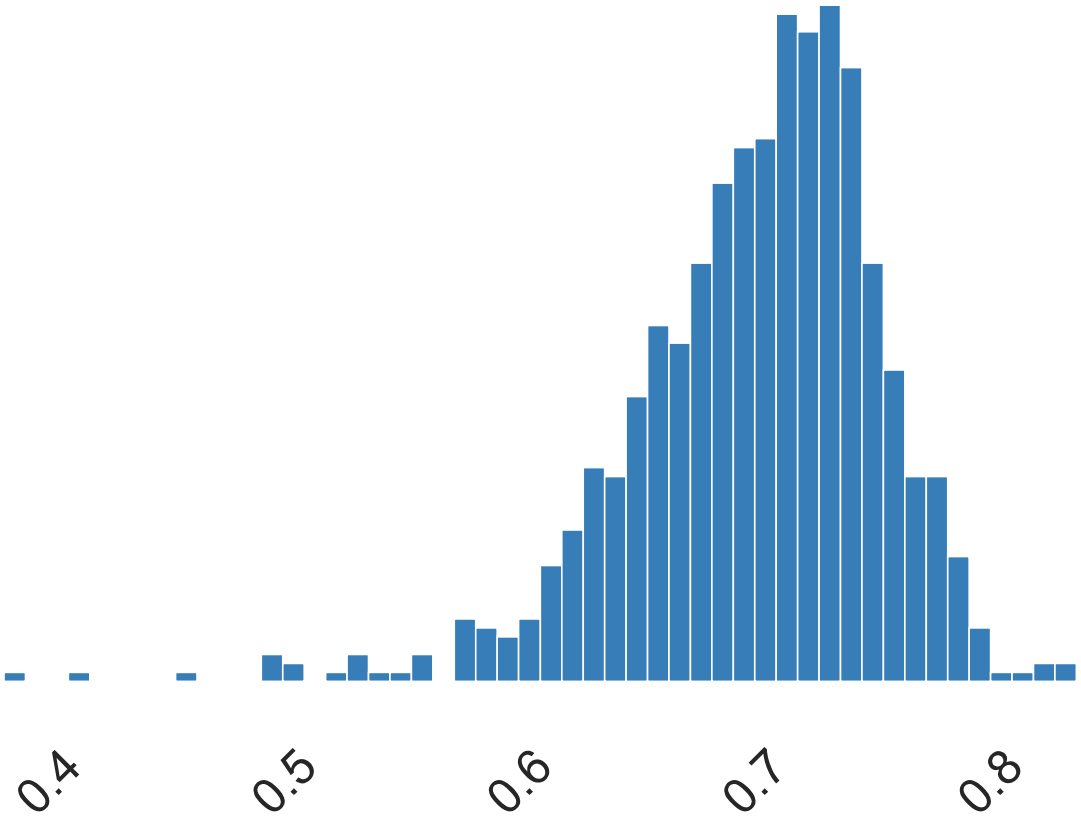
Value	Count	Frequency (%)
278217	1	0.1%
239093	1	0.1%
229195	1	0.1%
228259	1	0.1%
227170	1	0.1%
225916	1	0.1%
225592	1	0.1%
221527	1	0.1%
221396	1	0.1%
219952	1	0.1%

Extent

Real number (ℝ)

Distinct	900
Distinct (%)	100.0%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Mean	0.69950793

Minimum	0.37985611
Maximum	0.83545455
Zeros	0
Zeros (%)	0.0%
Negative	0
Negative (%)	0.0%
Memory size	7.2 KiB



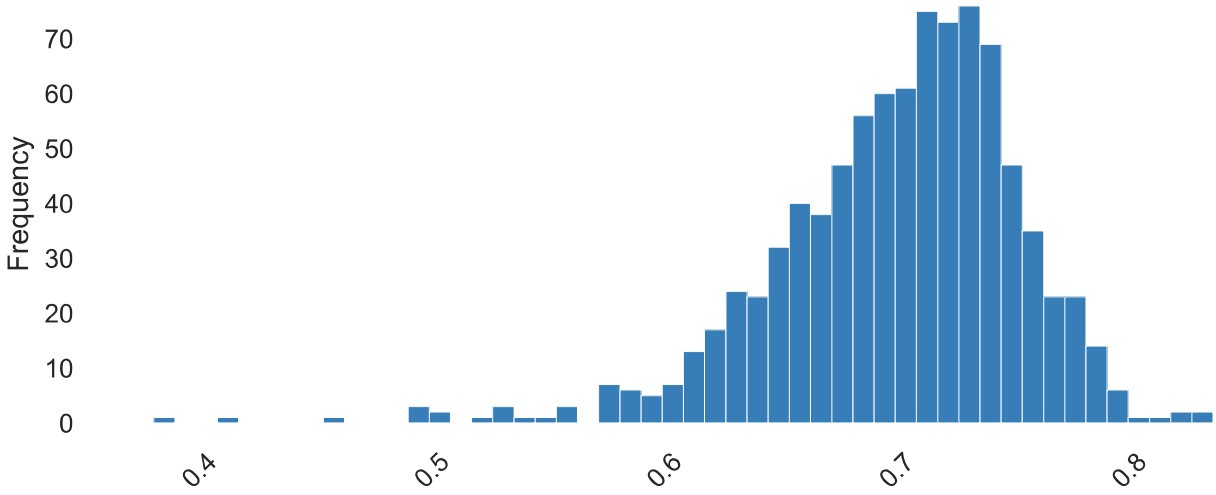
Quantile statistics

Minimum	0.37985611
5-th percentile	0.61168007
Q1	0.6708691
median	0.70736696
Q3	0.73499142
95-th percentile	0.77236549
Maximum	0.83545455
Range	0.45559843
Interquartile range (IQR)	0.06412232

Descriptive statistics

Standard deviation	0.0534682

Coefficient of variation (CV)	0.076436876
Kurtosis	3.3413837
Mean	0.69950793
Median Absolute Deviation (MAD)	0.03087709
Skewness	-1.1515048
Sum	629.55713
Variance	0.0028588484
Monotonicity	Not monotonic



Histogram with fixed size bins (bins=50)

Value	Count	Frequency (%)
0.758650579	1	0.1%
0.675596941	1	0.1%
0.728998849	1	0.1%
0.679589556	1	0.1%
0.714973571	1	0.1%
0.734356566	1	0.1%
0.744780007	1	0.1%
0.696992892	1	0.1%
0.594050406	1	0.1%
0.547433097	1	0.1%
Other values (890)	890	98.9%

Value	Count	Frequency (%)
0.379856115	1	0.1%
0.414153748	1	0.1%
0.454188929	1	0.1%
0.491002009	1	0.1%
0.491460095	1	0.1%
0.496936979	1	0.1%
0.500546035	1	0.1%
0.507352075	1	0.1%
0.522278775	1	0.1%
0.526259635	1	0.1%

Value	Count	Frequency (%)
0.835454545	1	0.1%
0.830632225	1	0.1%
0.824319225	1	0.1%
0.817388614	1	0.1%
0.812265444	1	0.1%
0.800255194	1	0.1%
0.796730297	1	0.1%
0.794371379	1	0.1%
0.792998657	1	0.1%
0.792771926	1	0.1%

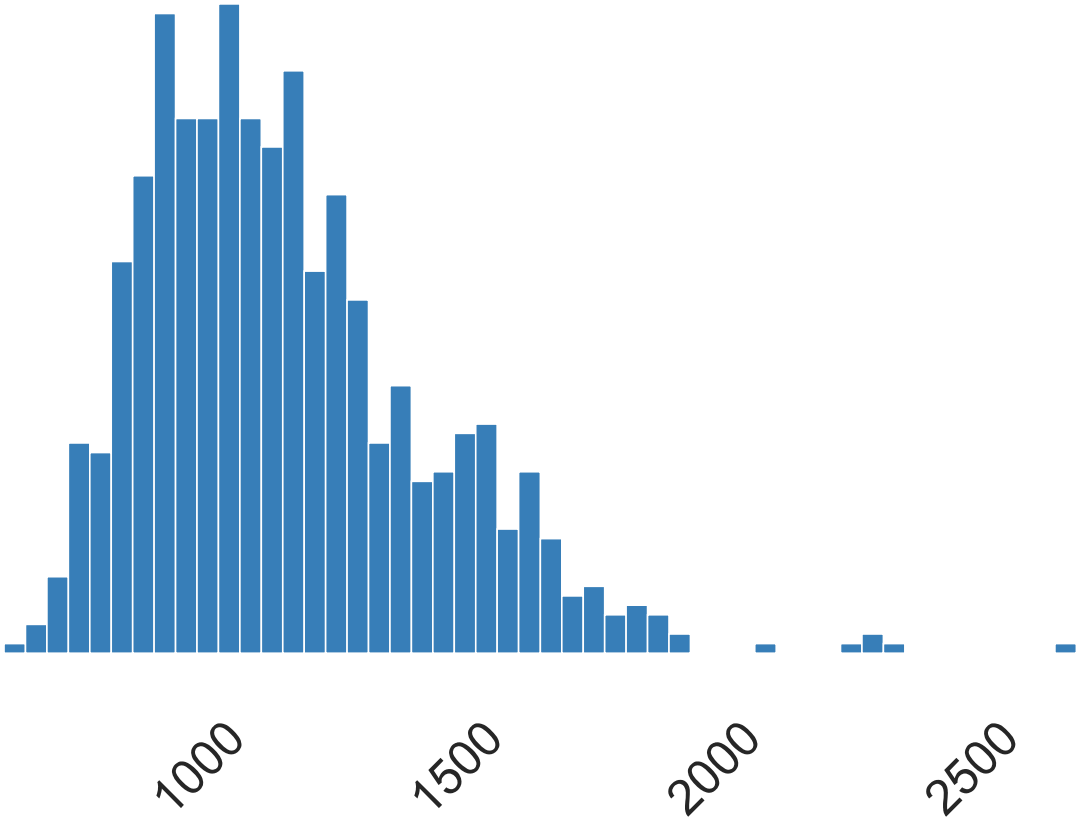
Perimeter

Real number (ℝ)

HIGH CORRELATION (This variable has a high overall correlation with 6 fields: Area, MajorAxisLength, MinorAxisLength, Eccentricity, ConvexArea, Class) UNIQUE

Distinct	900
Distinct (%)	100.0%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Mean	1165.9066

Minimum	619.074
Maximum	2697.753
Zeros	0
Zeros (%)	0.0%
Negative	0
Negative (%)	0.0%
Memory size	7.2 KiB



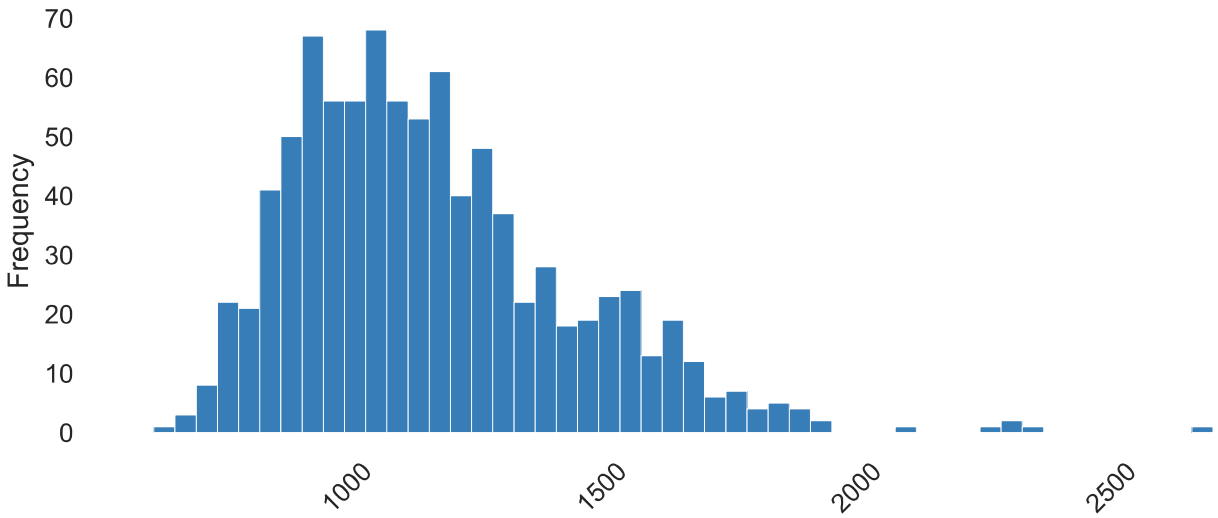
Quantile statistics

Minimum	619.074
5-th percentile	804.51845
Q1	966.41075
median	1119.509
Q3	1308.3898
95-th percentile	1661.5969
Maximum	2697.753
Range	2078.679
Interquartile range (IQR)	341.979

Descriptive statistics

Standard deviation	273.76432

Coefficient of variation (CV)	0.23480809
Kurtosis	1.7447061
Mean	1165.9066
Median Absolute Deviation (MAD)	172.495
Skewness	1.0177611
Sum	1049316
Variance	74946.9
Monotonicity	Not monotonic



Histogram with fixed size bins (bins=50)

Value	Count	Frequency (%)
1184.04	1	0.1%
1394.1	1	0.1%
1388.684	1	0.1%
1502.661	1	0.1%
1219.105	1	0.1%
950.297	1	0.1%
1290.239	1	0.1%
1144.036	1	0.1%
1570.502	1	0.1%
1295.377	1	0.1%
Other values (890)	890	98.9%

Value	Count	Frequency (%)
619.074	1	0.1%
678.815	1	0.1%
683.004	1	0.1%
699.415	1	0.1%
713.775	1	0.1%
713.94	1	0.1%
718.847	1	0.1%
719.935	1	0.1%
727.561	1	0.1%
734.102	1	0.1%

Value	Count	Frequency (%)
2697.753	1	0.1%
2352.029	1	0.1%
2303.69	1	0.1%
2289.889	1	0.1%
2253.557	1	0.1%
2098.263	1	0.1%
1947.46	1	0.1%
1942.05	1	0.1%
1893.414	1	0.1%
1876.307	1	0.1%

Class

Categorical

HIGH CORRELATION (This variable has a high overall correlation with 4 fields: Area, MajorAxisLength, ConvexArea, Perimeter) UNIFORM

Distinct	2
Distinct (%)	0.2%
Missing	0
Missing (%)	0.0%
Memory size	7.2 KiB

Length

Max length	7
Median length	6
Mean length	6
Min length	5

Characters and Unicode

Total characters	5400	
Distinct characters	8	
Distinct categories	2 (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%	

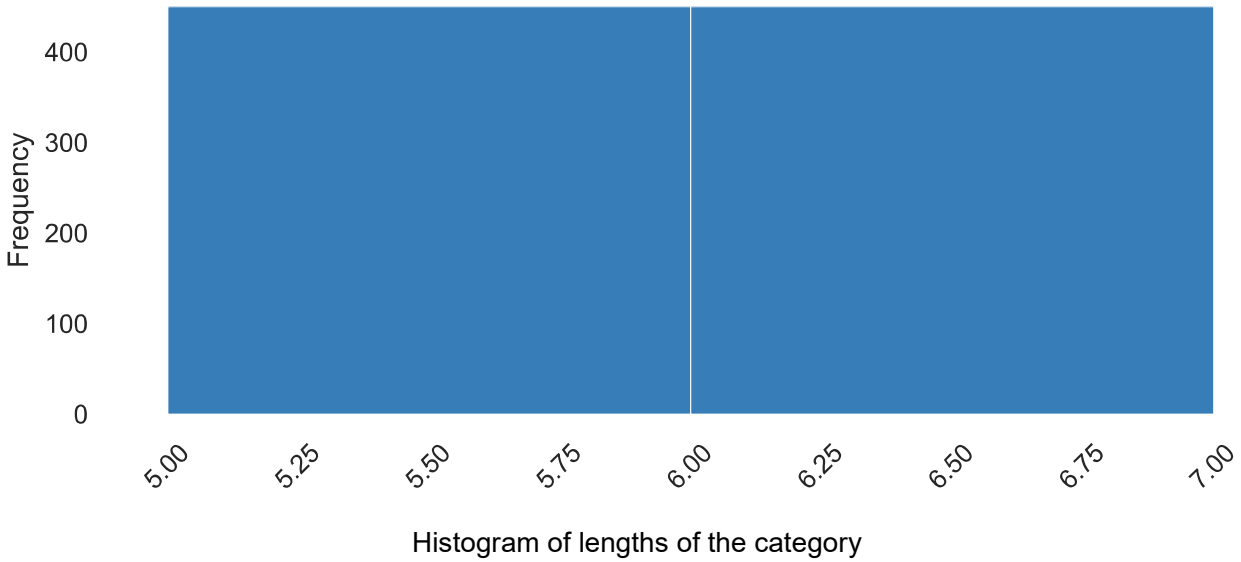
Sample

1st row	Kecimen
2nd row	Kecimen
3rd row	Kecimen
4th row	Kecimen
5th row	Kecimen

Common Values

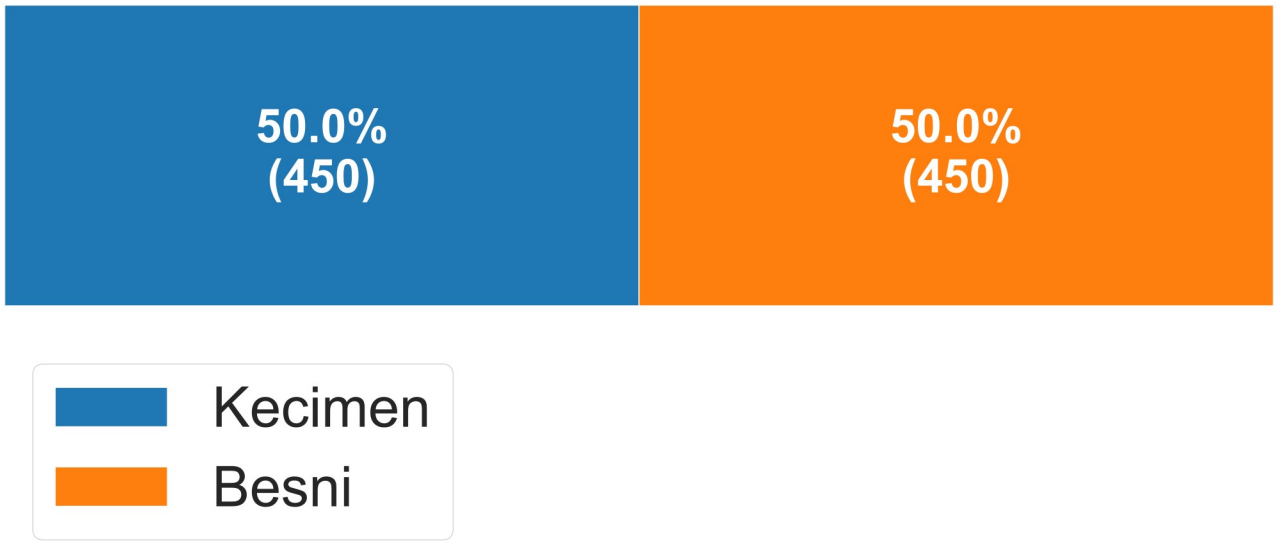
Value	Count	Frequency (%)
Kecimen	450	50.0%
Besni	450	50.0%

Length



Histogram of lengths of the category

Common Values (Plot)



Value	Count	Frequency (%)
kecimen	450	50.0%
besni	450	50.0%

Most occurring characters

Value	Count	Frequency (%)
e	1350	25.0%
i	900	16.7%
n	900	16.7%
K	450	8.3%
c	450	8.3%
m	450	8.3%
B	450	8.3%
s	450	8.3%

Most occurring categories

Value	Count	Frequency (%)
Lowercase Letter	4500	83.3%
Uppercase Letter	900	16.7%

Most frequent character per category

Lowercase Letter

Value	Count	Frequency (%)
e	1350	30.0%
i	900	20.0%
n	900	20.0%
c	450	10.0%
m	450	10.0%
s	450	10.0%

Uppercase Letter

Value	Count	Frequency (%)
K	450	50.0%
B	450	50.0%

Most occurring scripts

Value	Count	Frequency (%)
Latin	5400	100.0%

Most frequent character per script

Latin

Value	Count	Frequency (%)
e	1350	25.0%
i	900	16.7%
n	900	16.7%
K	450	8.3%
c	450	8.3%
m	450	8.3%
B	450	8.3%
s	450	8.3%

Most occurring blocks

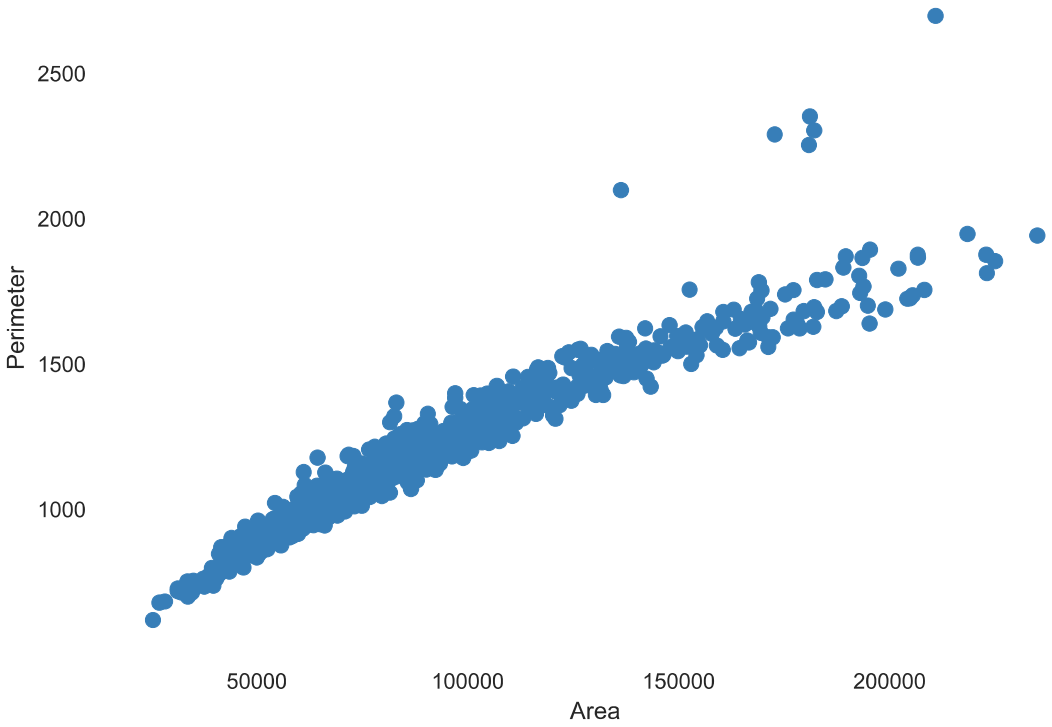
Value	Count	Frequency (%)
ASCII	5400	100.0%

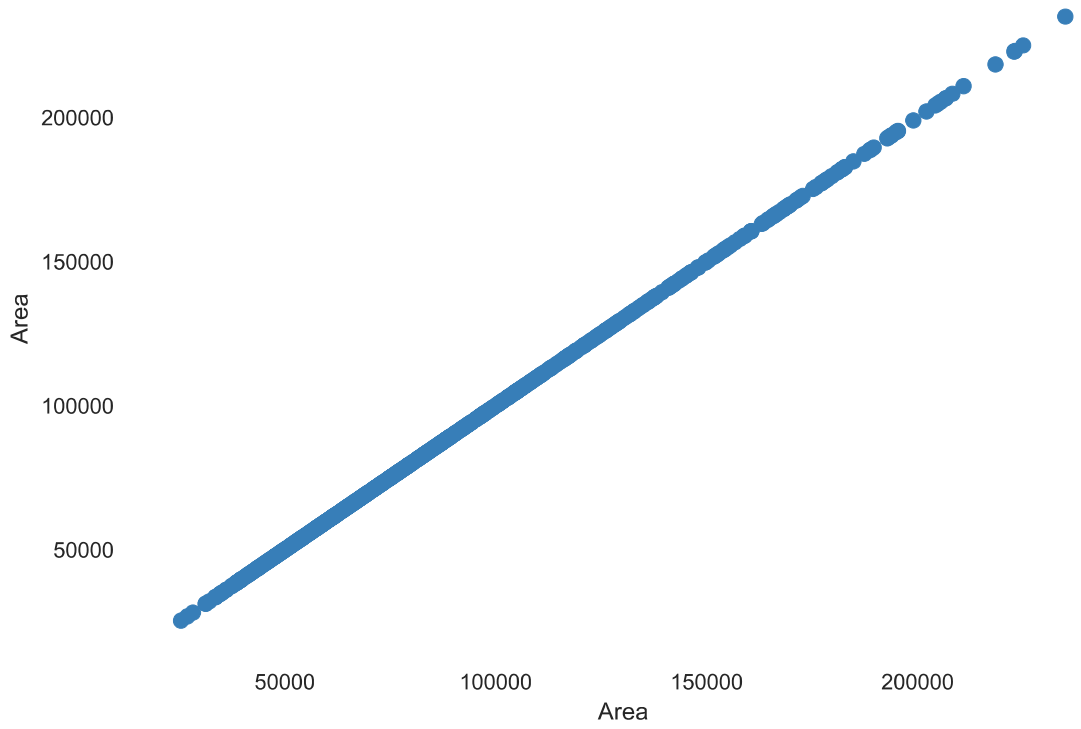
Most frequent character per block

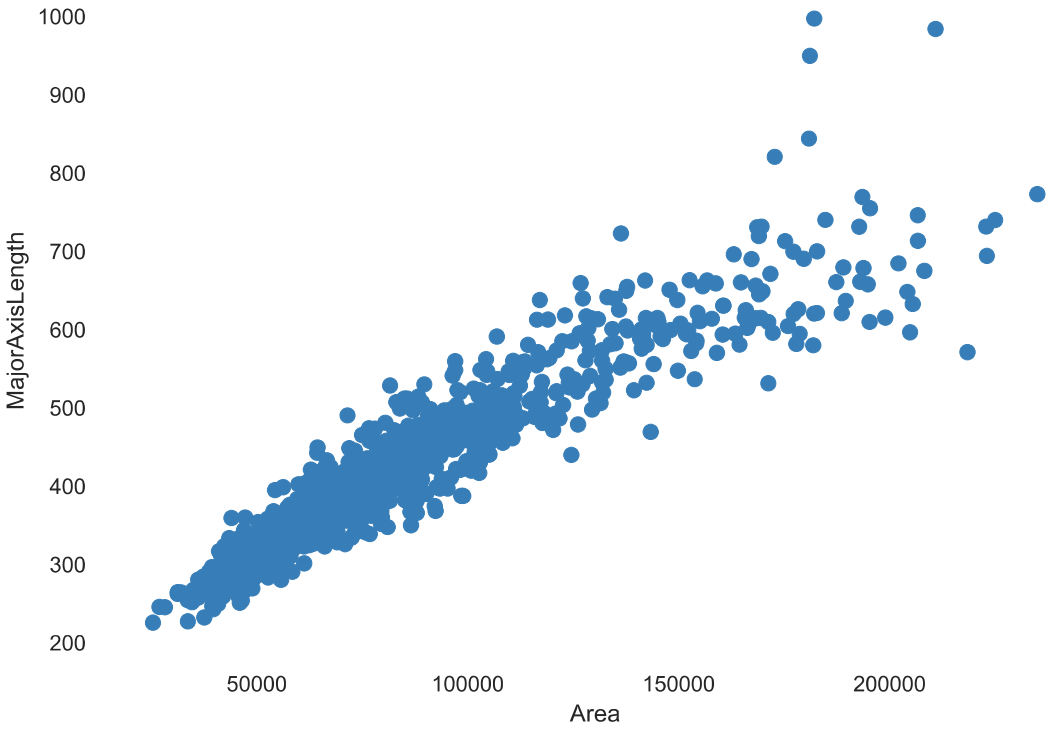
ASCII

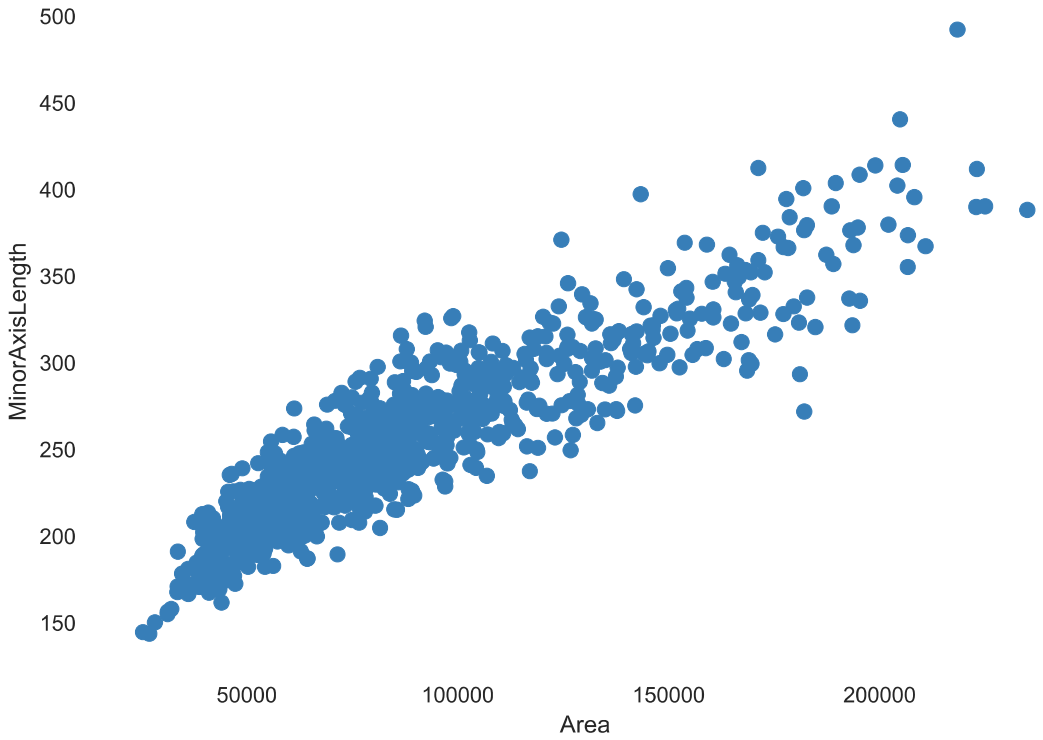
Value	Count	Frequency (%)
e	1350	25.0%
i	900	16.7%
n	900	16.7%
K	450	8.3%
c	450	8.3%
m	450	8.3%
B	450	8.3%
s	450	8.3%

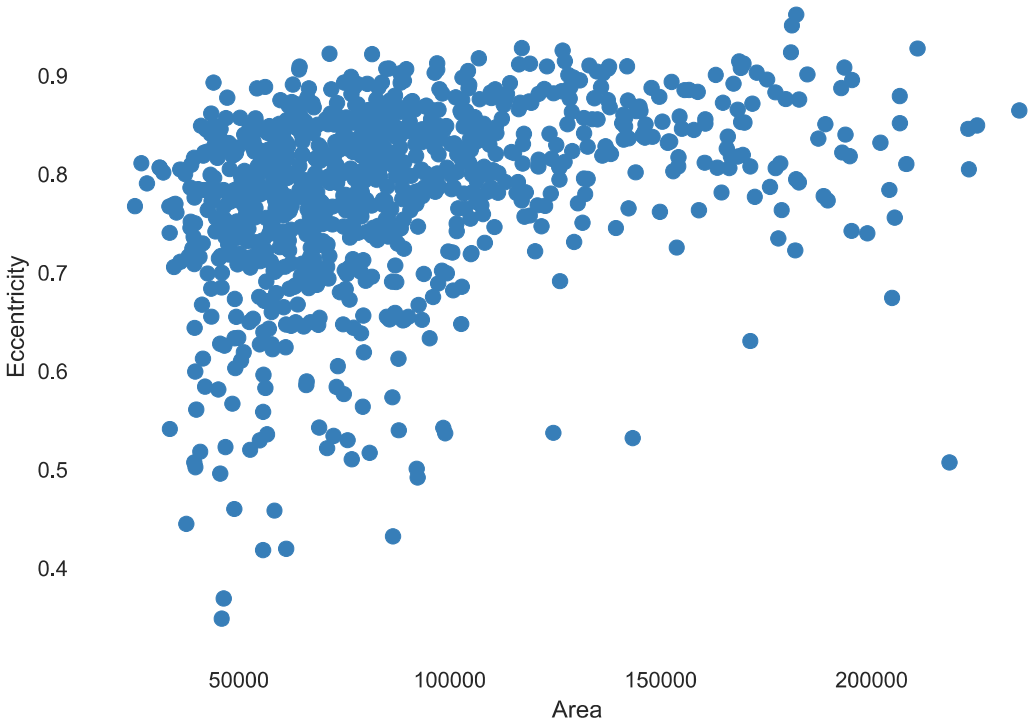
Interactions

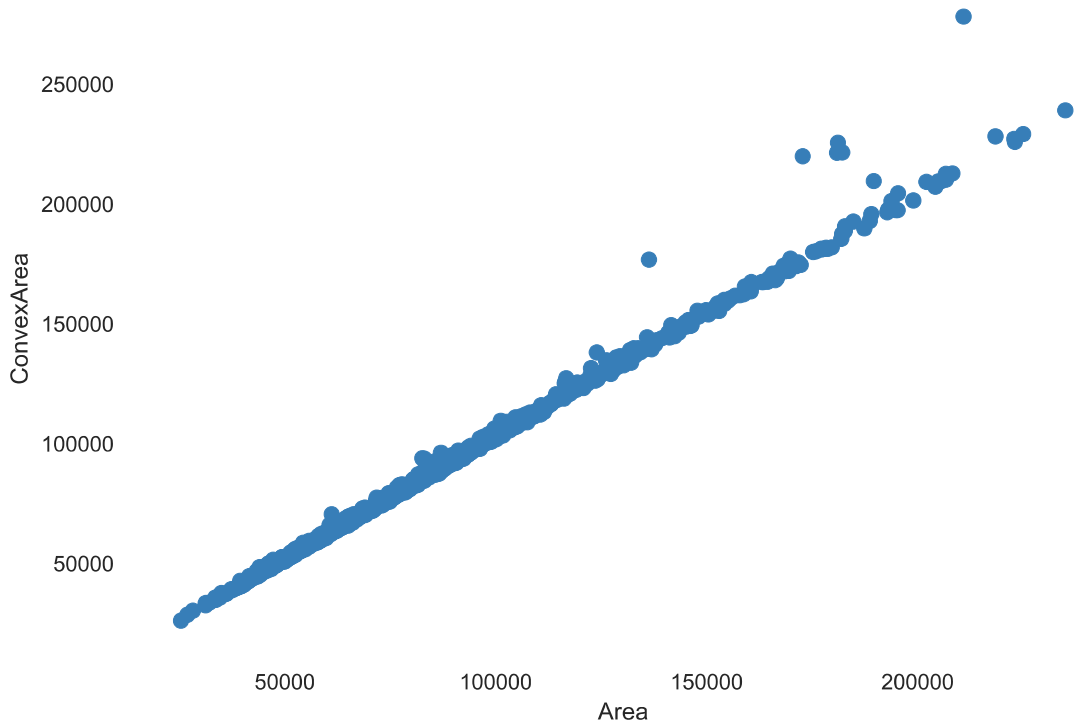


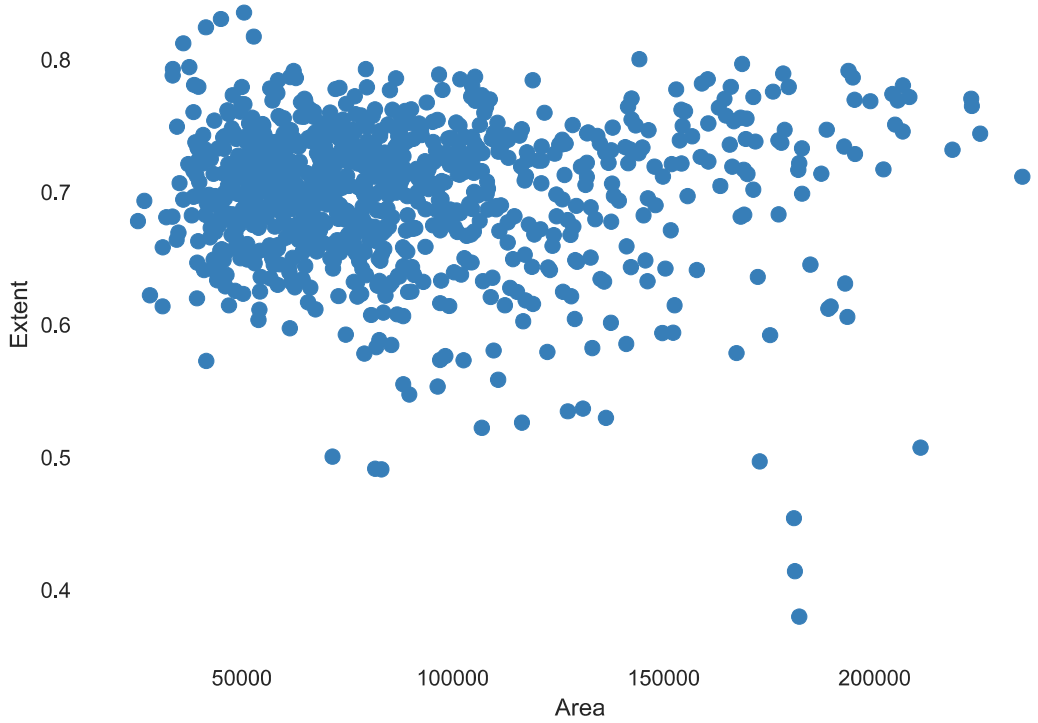

200000
150000
100000
50000
Area
50000
100000
150000
200000
Area

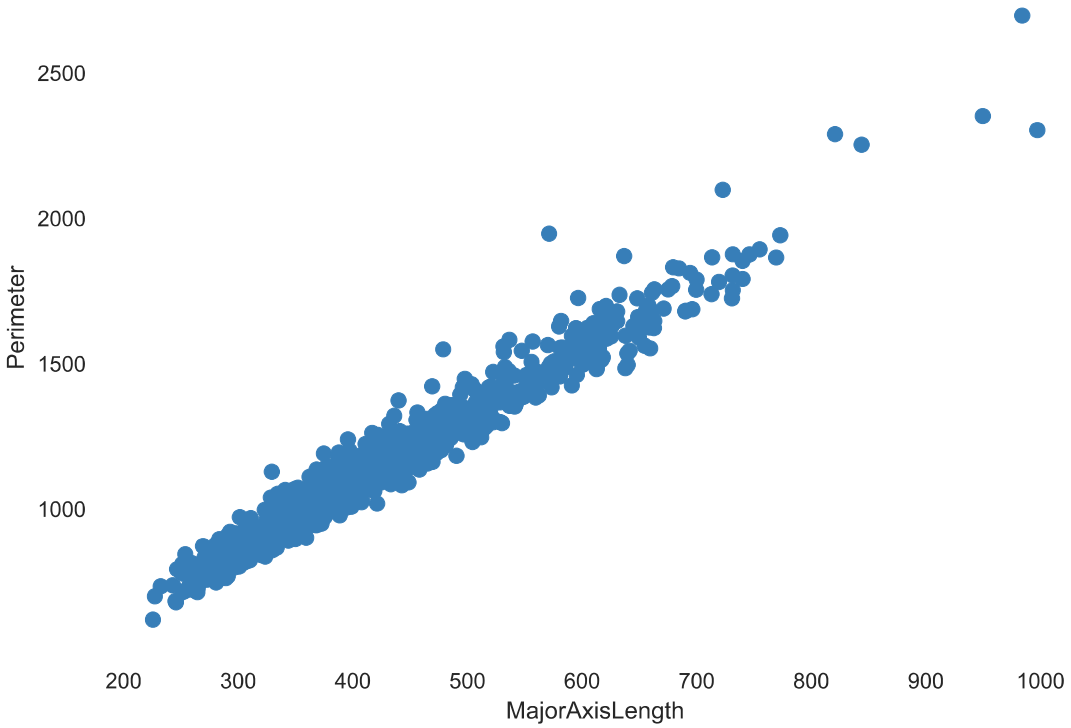

1000
900
800
700
600
500
400
300
200
MajorAxisLength
50000
100000
150000
200000
Area

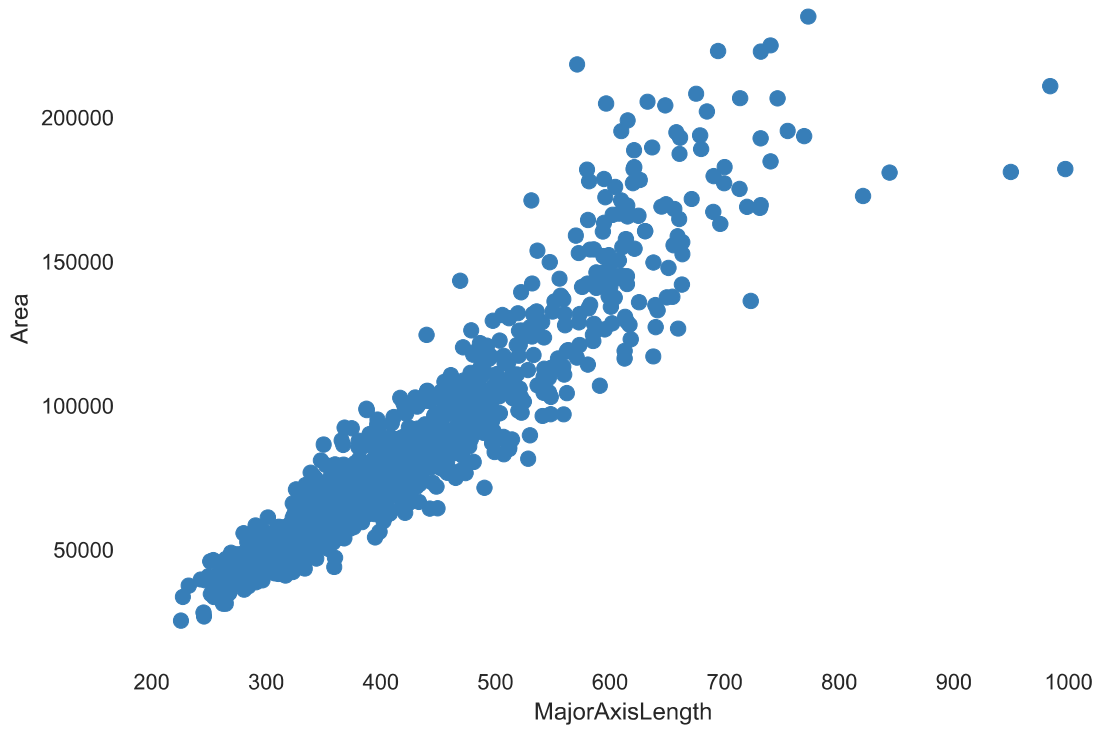

500
450
400
350
300
250
200
150
MinorAxisLength
50000
100000
150000
200000
Area

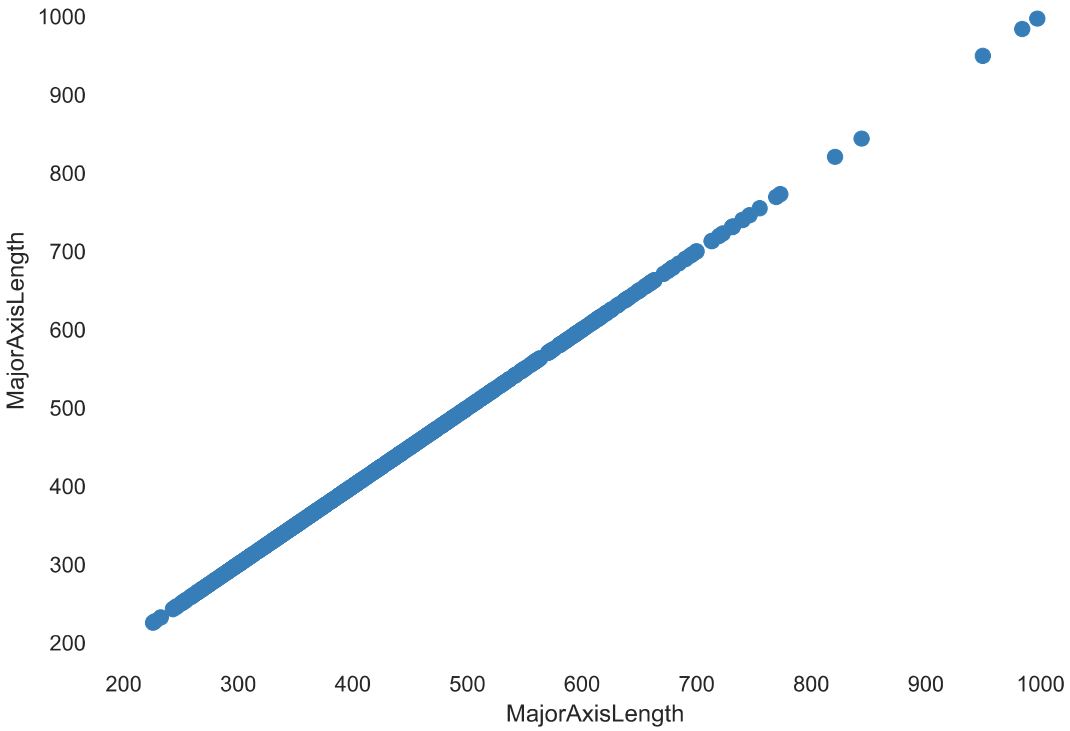

0.9
0.8
0.7
0.6
0.5
0.4
Eccentricity
50000
100000
150000
200000
Area

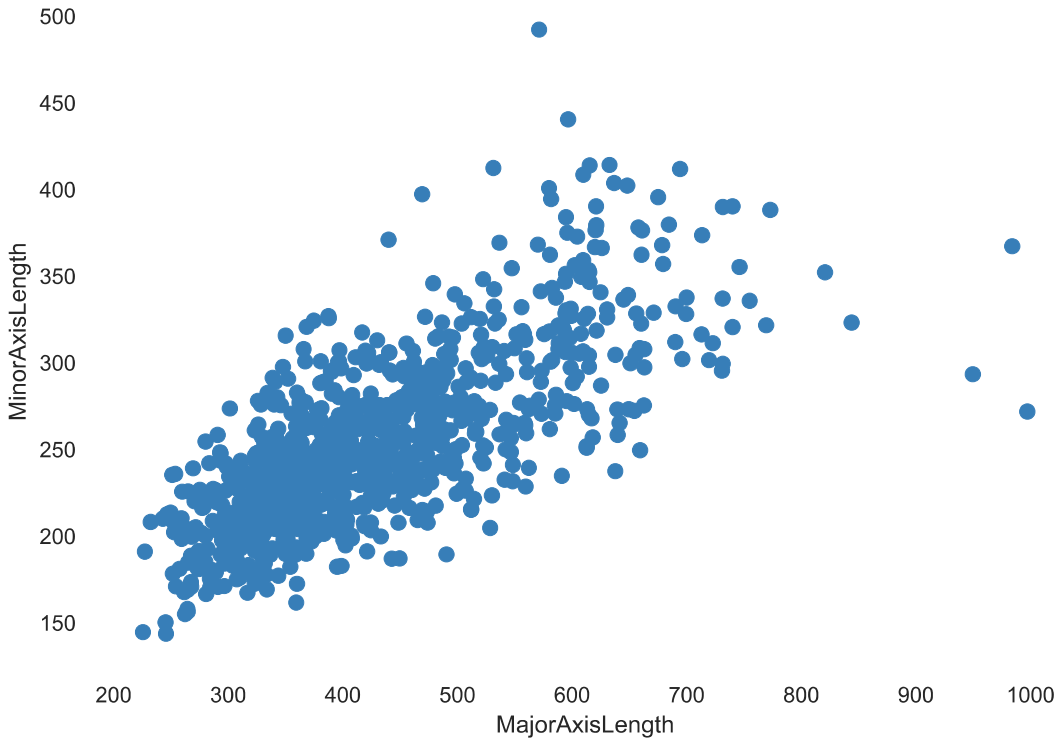

250000
200000
150000
100000
50000
ConvexArea
50000
100000
150000
200000
Area

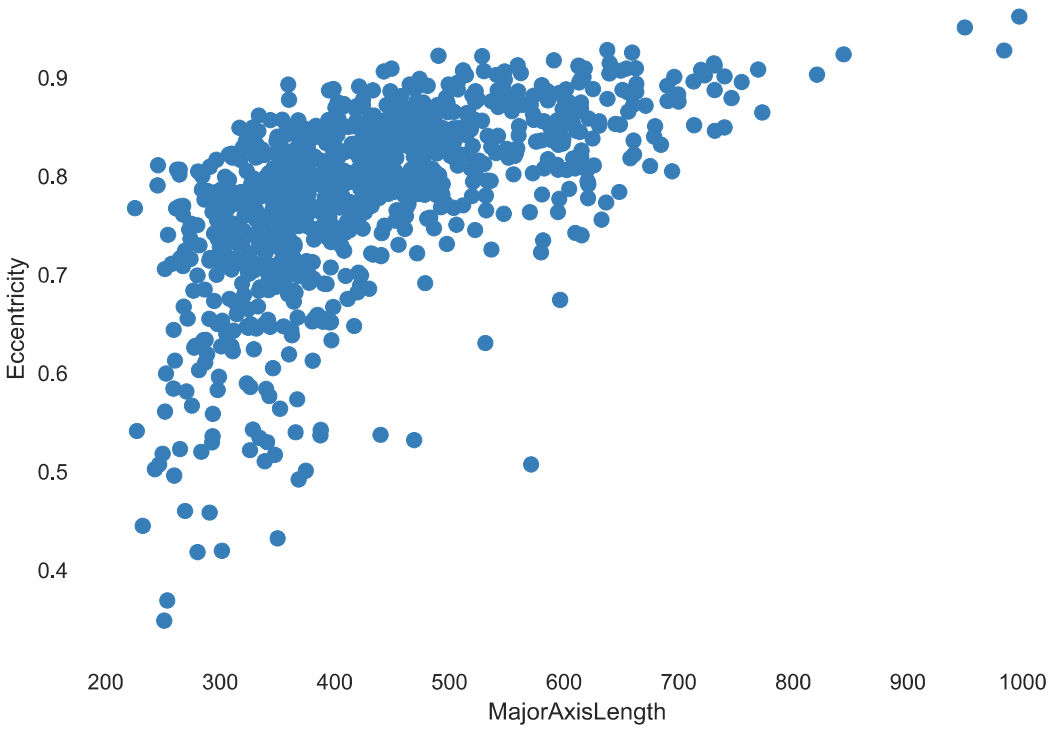

0.8
0.7
0.6
0.5
0.4
Extent
50000
100000
150000
200000
Area

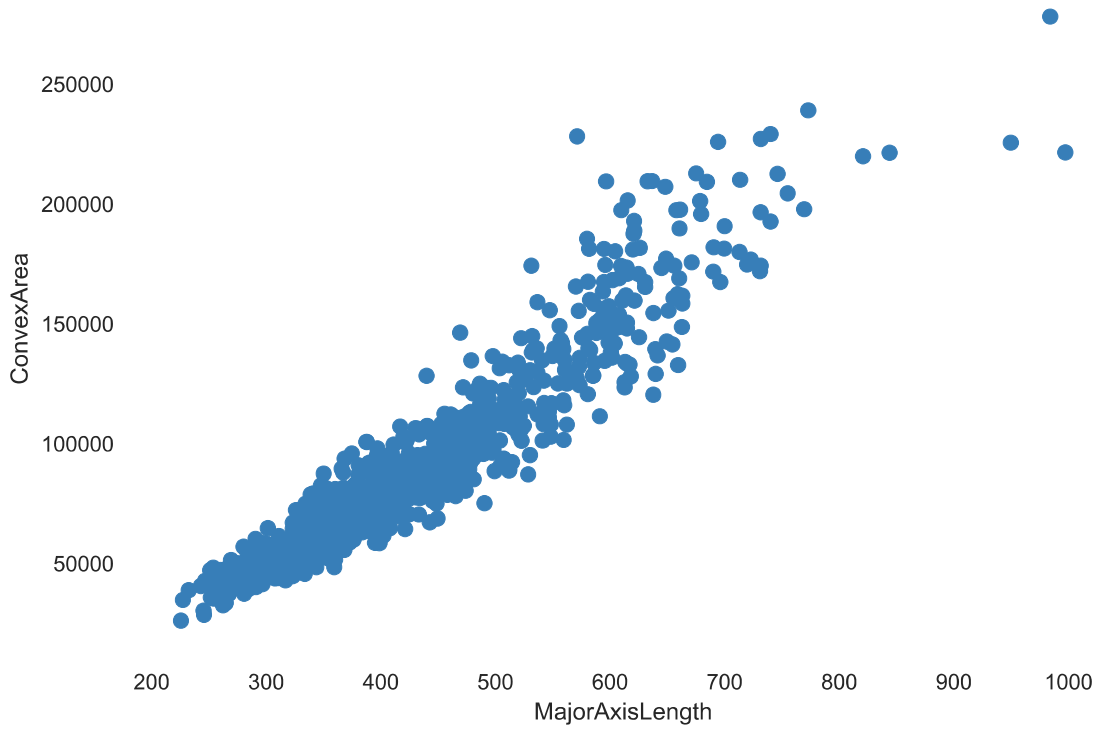

2500
2000
1500
1000
Perimeter
200
300
400
500
600
700
800
900
1000
MajorAxisLength

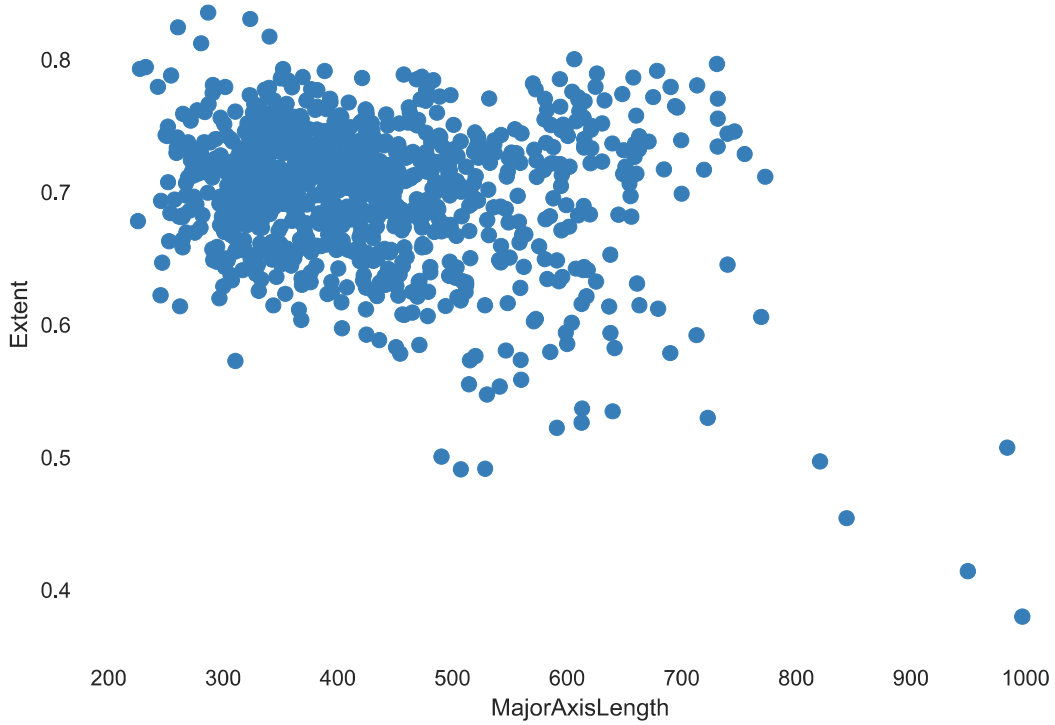

200000
150000
100000
50000
Area
200
300
400
500
600
700
800
900
1000
MajorAxisLength

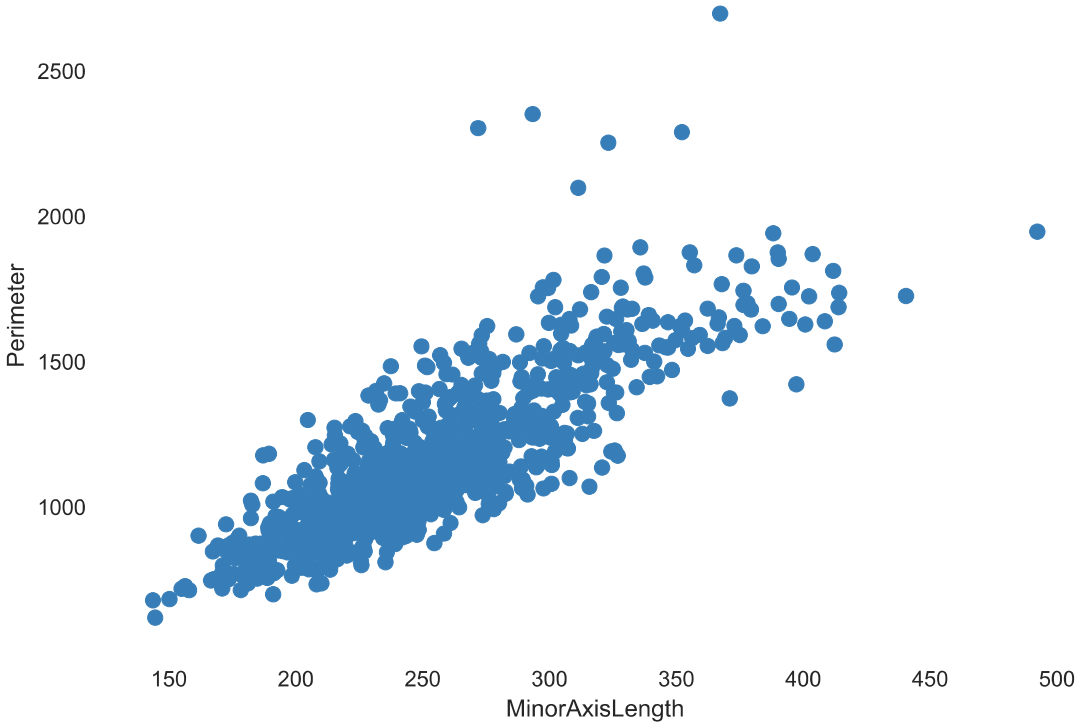

1000
900
800
700
600
500
400
300
200
MajorAxisLength
200
300
400
500
600
700
800
900
1000
MajorAxisLength

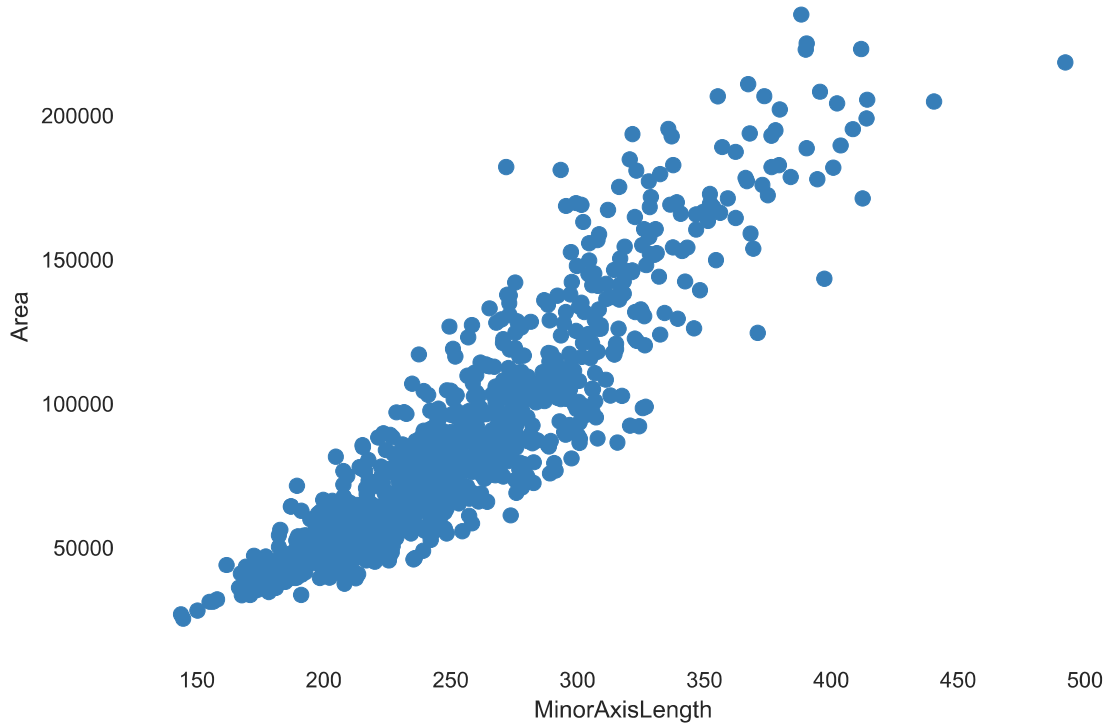

MinorAxisLength
500
450
400
350
300
250
200
150
200
300
400
500
600
700
800
900
1000
MajorAxisLength

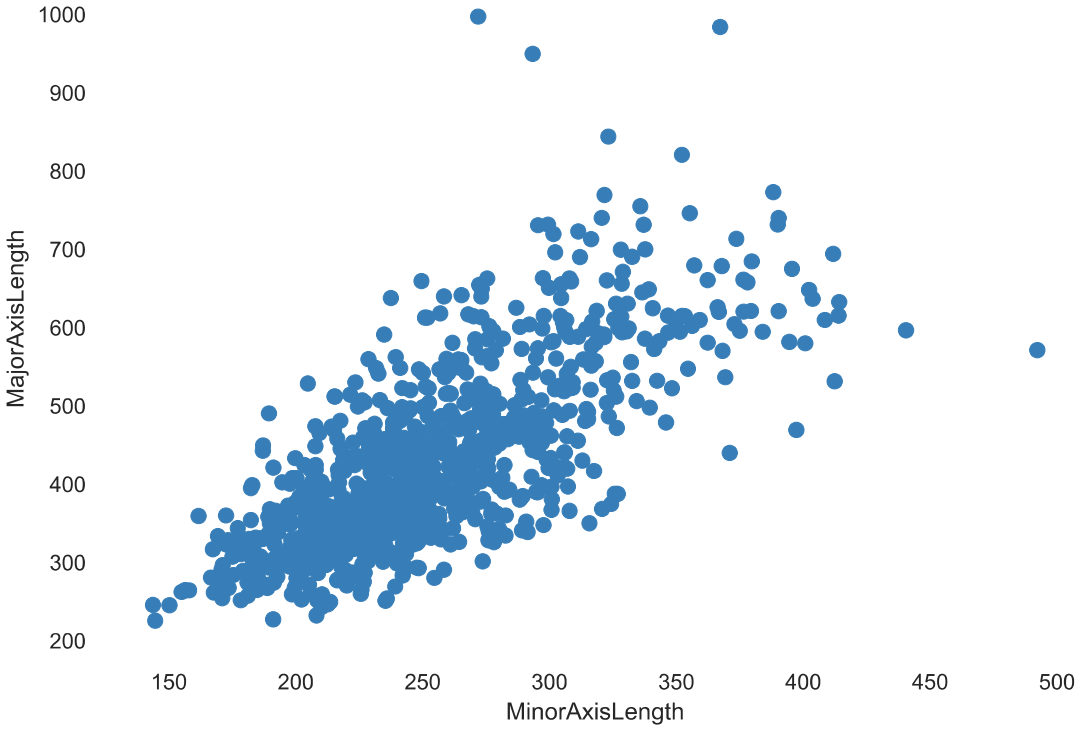

Eccentricity
0.9
0.8
0.7
0.6
0.5
0.4
200
300
400
500
600
700
800
900
1000
MajorAxisLength

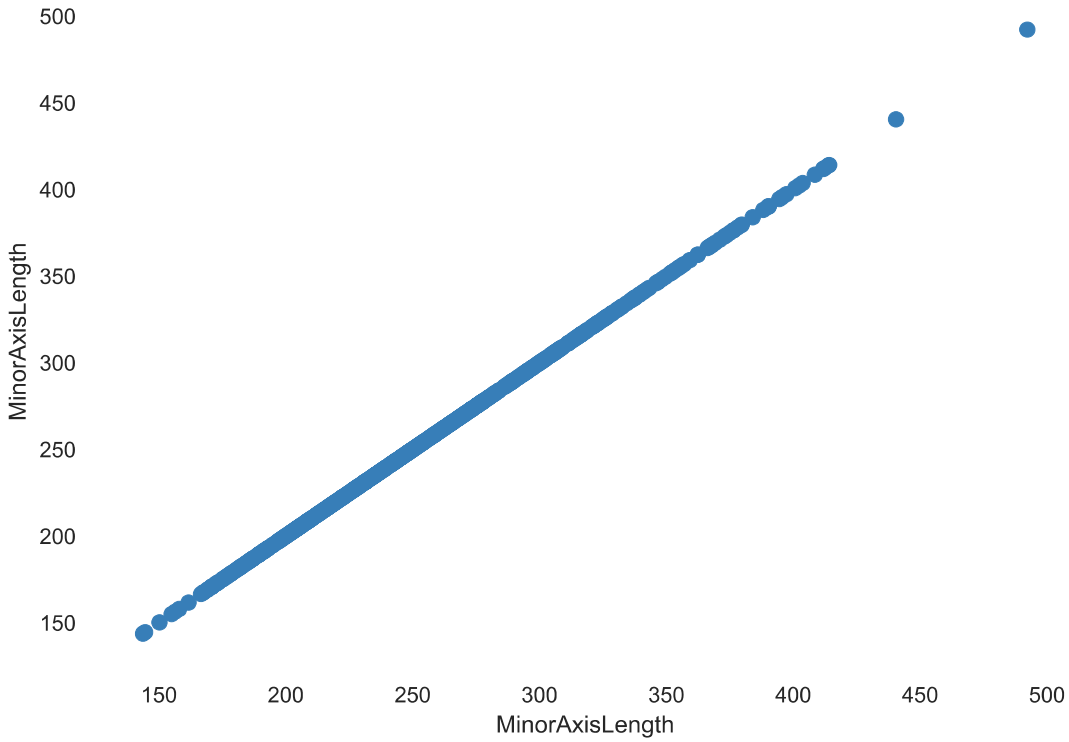

250000
200000
150000
100000
50000
ConvexArea
200
300
400
500
600
700
800
900
1000
MajorAxisLength

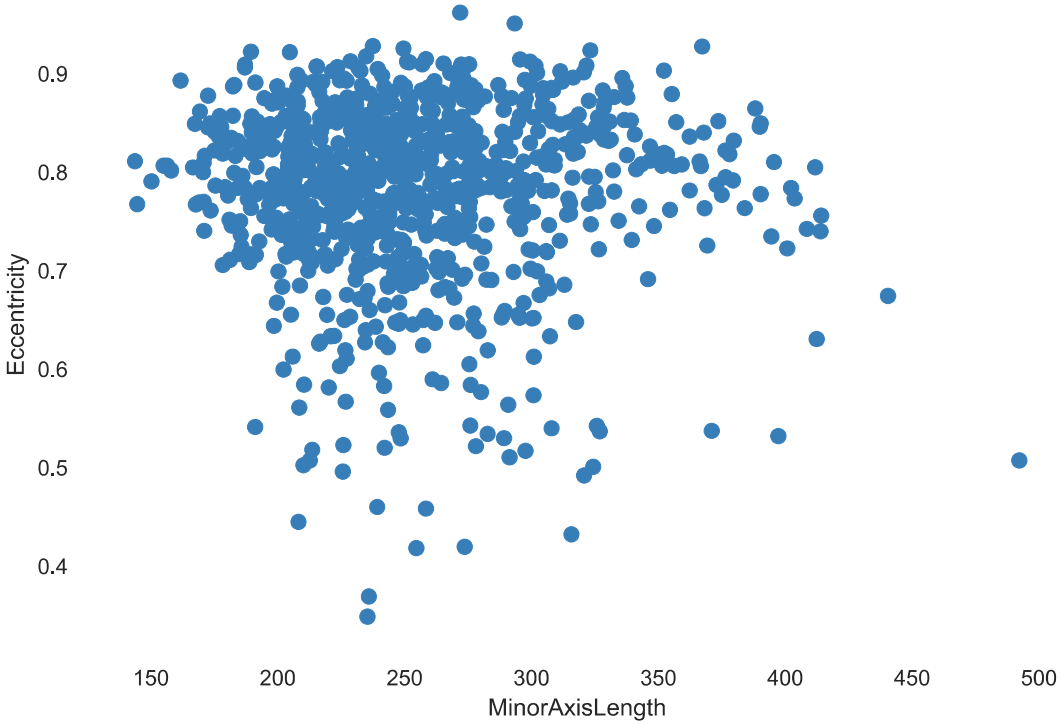

0.8
0.7
0.6
0.5
0.4
Extent
200
300
400
500
600
700
800
900
1000
MajorAxisLength

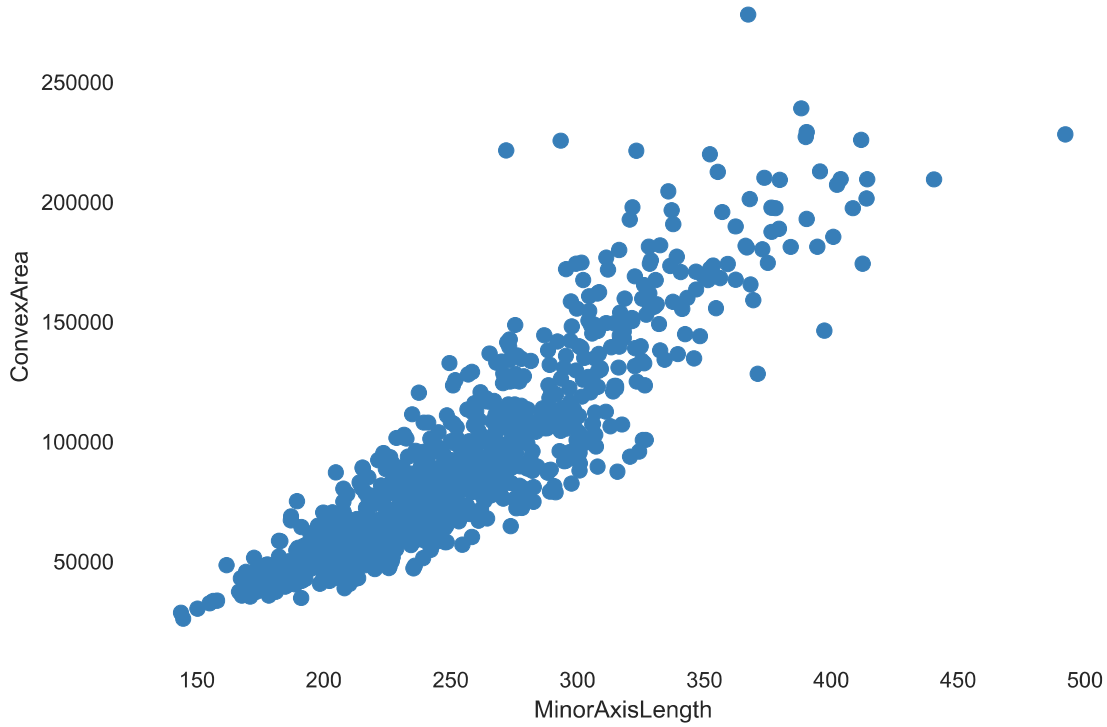

2500
2000
1500
1000
Perimeter
150
200
250
300
350
400
450
500
MinorAxisLength

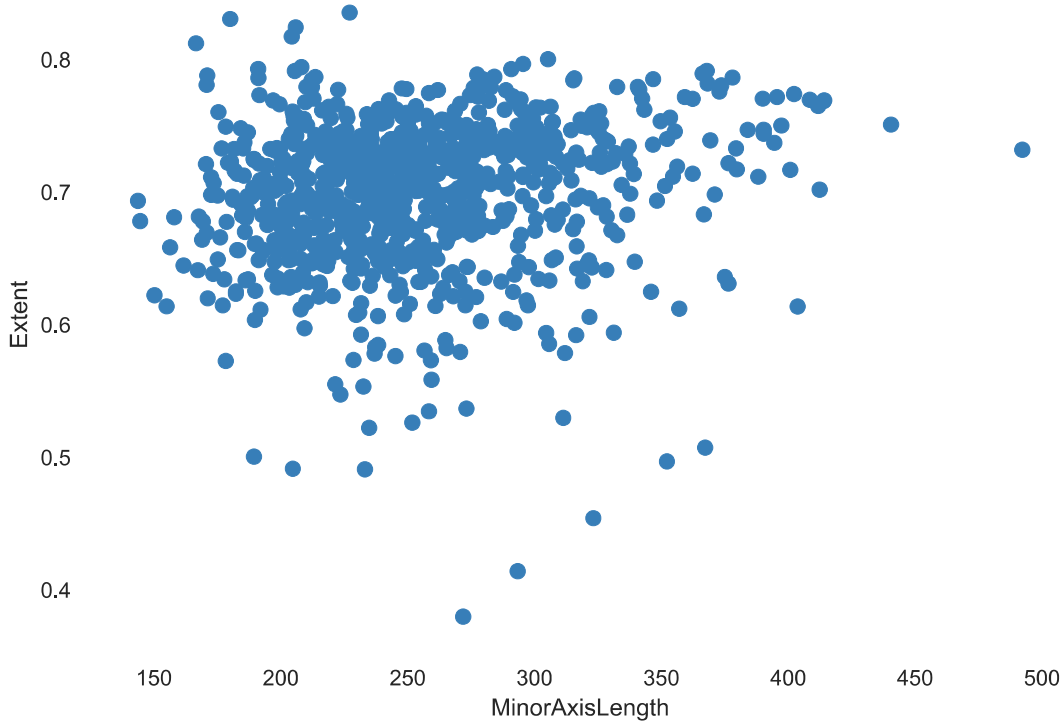

200000
150000
100000
50000
Area
150
200
250
300
350
400
450
500
MinorAxisLength

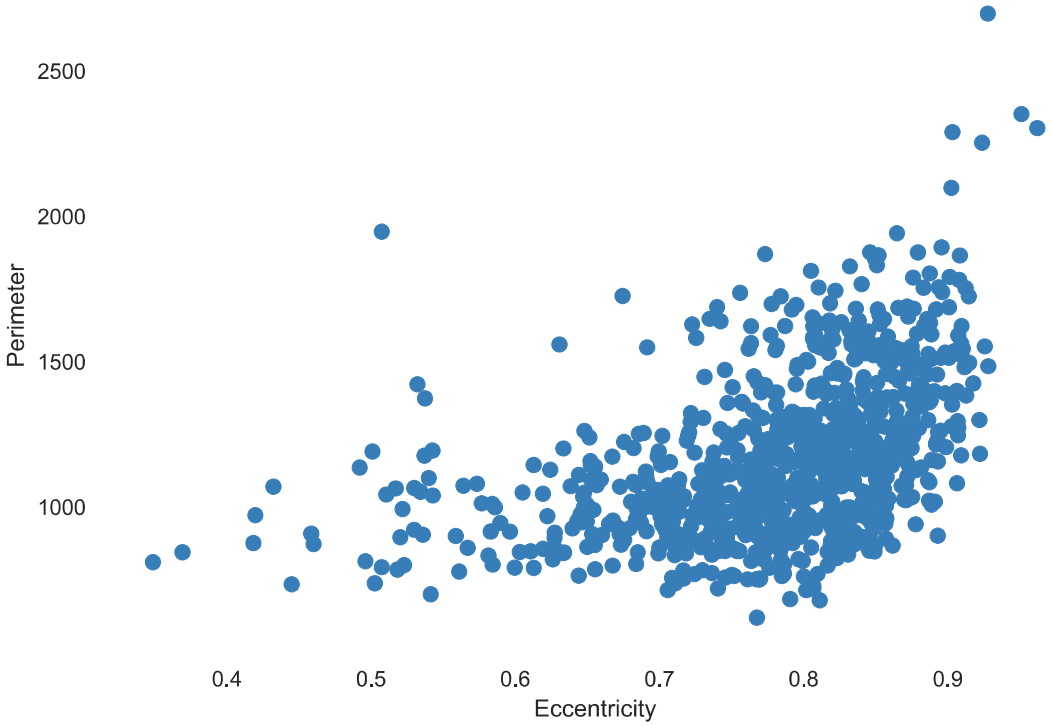

1000
900
800
700
600
500
400
300
200
MajorAxisLength
150
200
250
300
350
400
450
500
MinorAxisLength

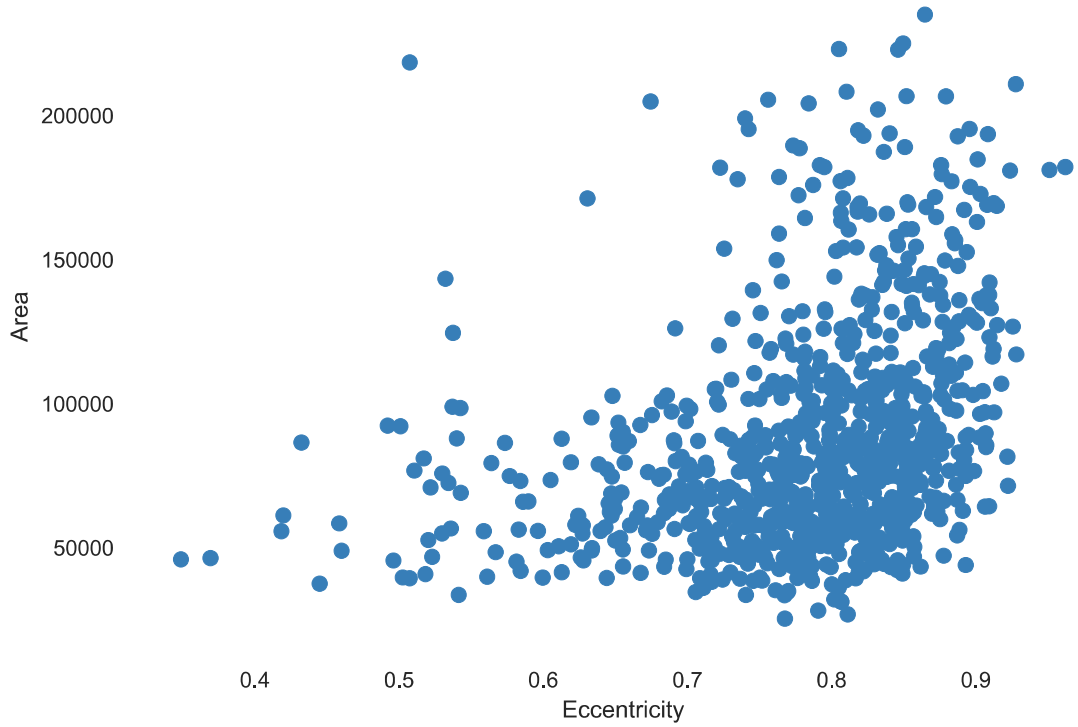

500
450
400
350
300
250
200
150
MinorAxisLength
150
200
250
300
350
400
450
500
MinorAxisLength

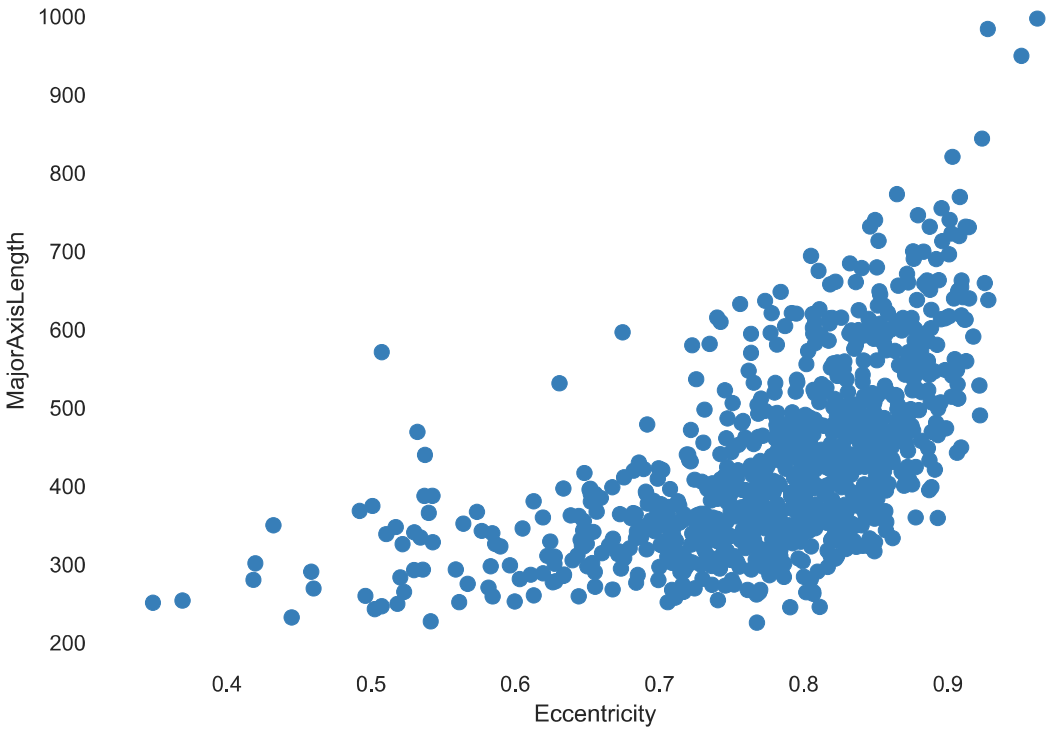

Eccentricity
0.9
0.8
0.7
0.6
0.5
0.4
150
200
250
300
350
400
450
500
MinorAxisLength

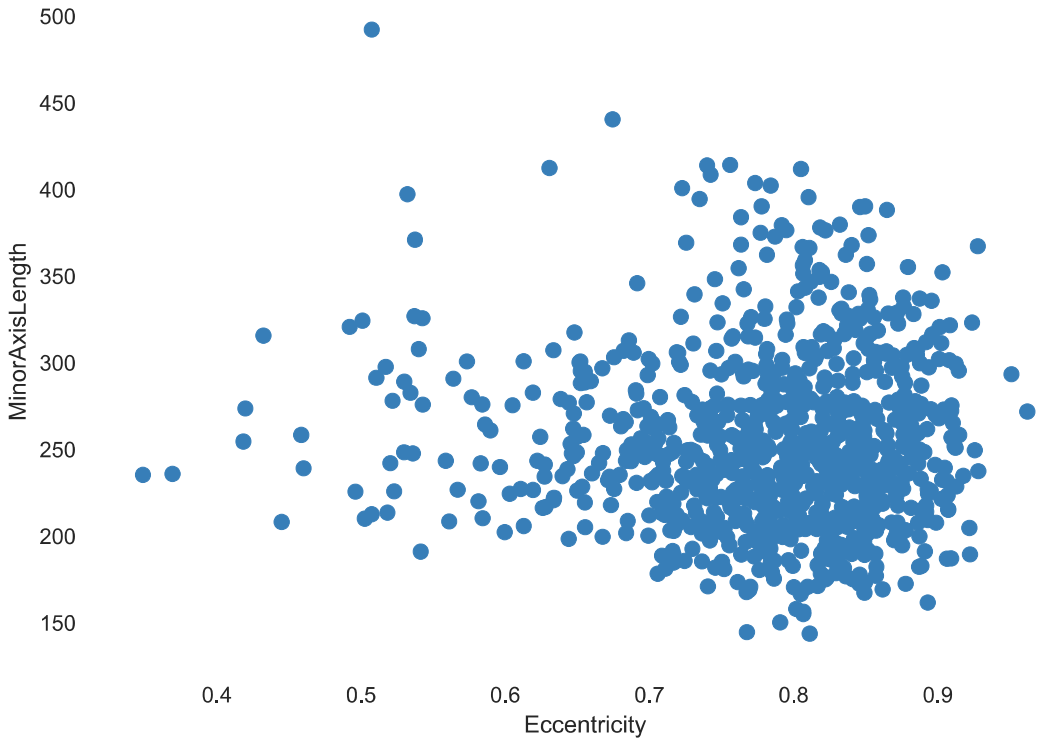

250000
200000
150000
100000
50000
ConvexArea
150
200
250
300
350
400
450
500
MinorAxisLength

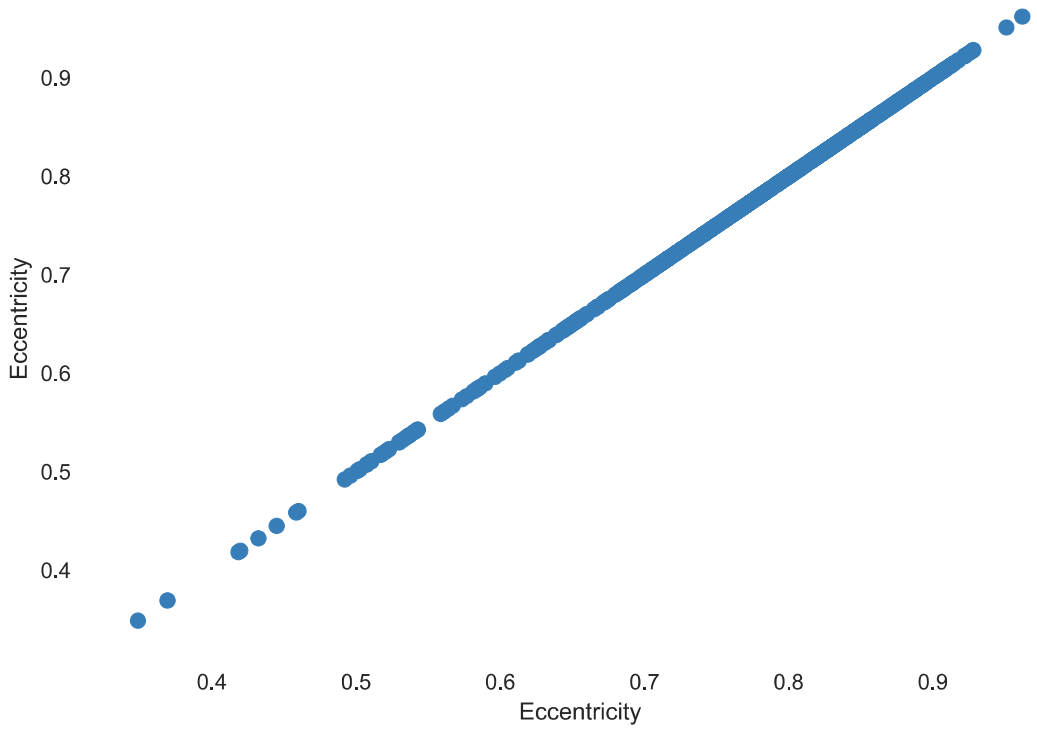

0.8
0.7
0.6
0.5
0.4
Extent
150
200
250
300
350
400
450
500
MinorAxisLength

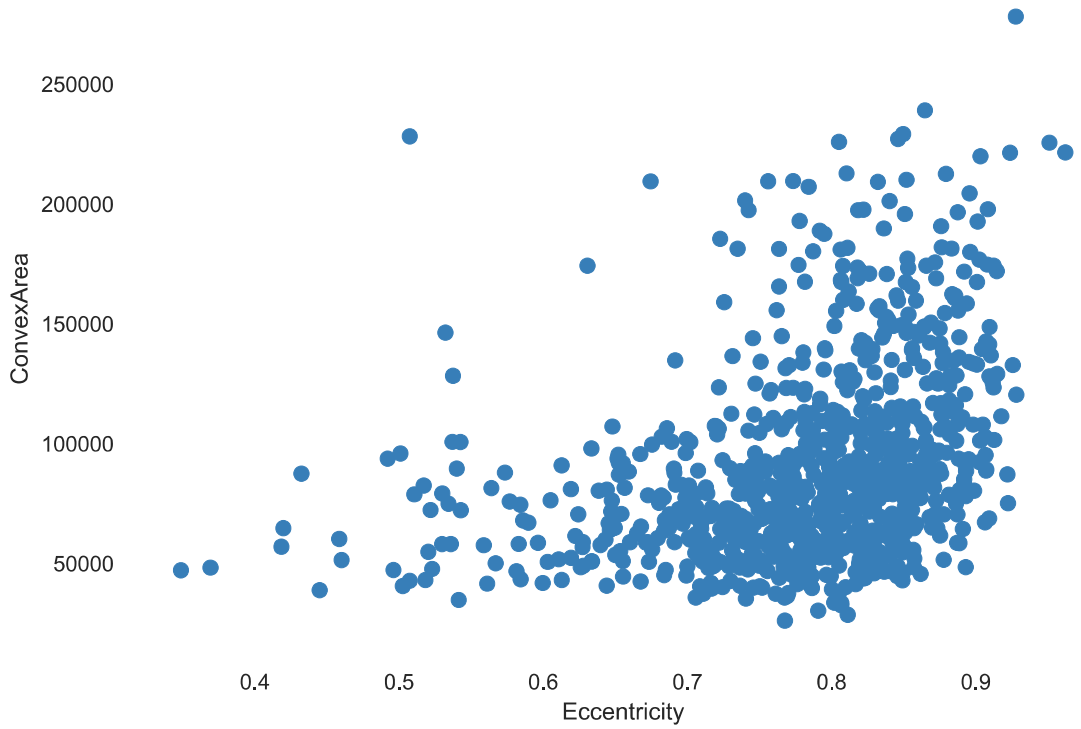

2500
2000
1500
1000
Perimeter
0.4
0.5
0.6
0.7
0.8
0.9
Eccentricity

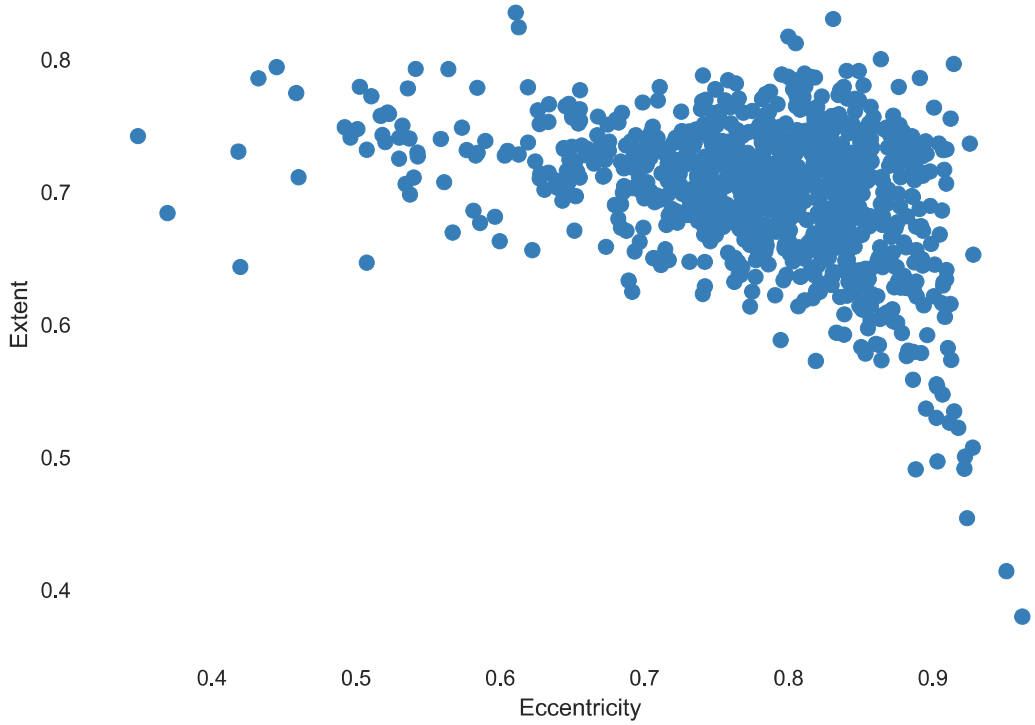

200000
150000
100000
50000
Area
0.4
0.5
0.6
0.7
0.8
0.9
Eccentricity

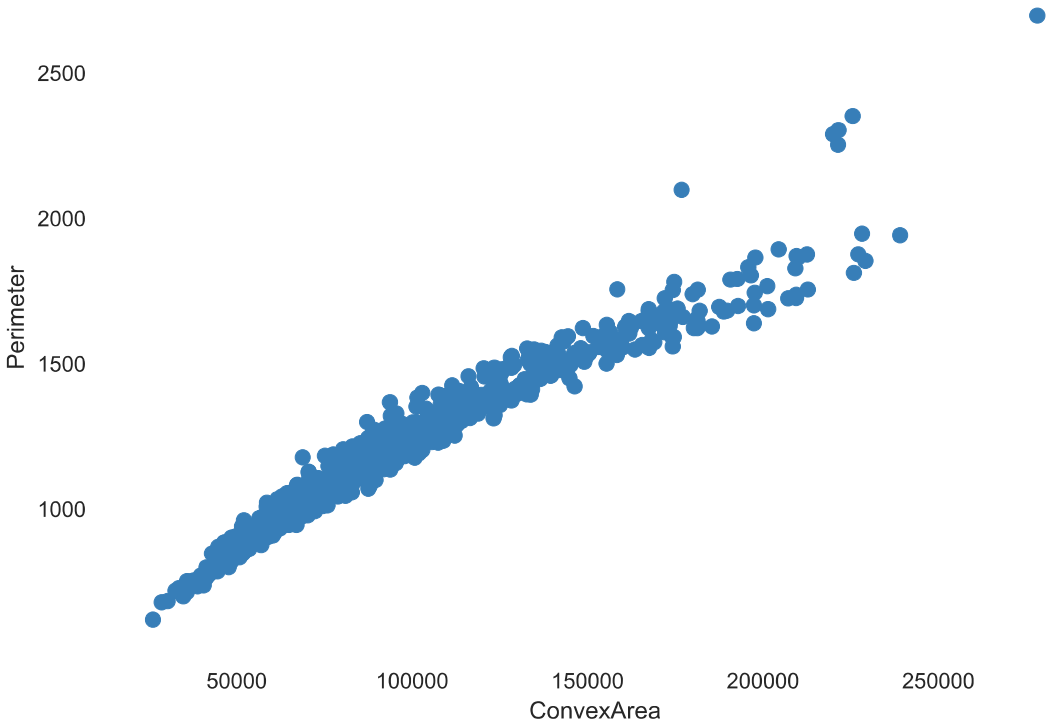

1000
900
800
700
600
500
400
300
200
MajorAxisLength
0.4
0.5
0.6
0.7
0.8
0.9
Eccentricity

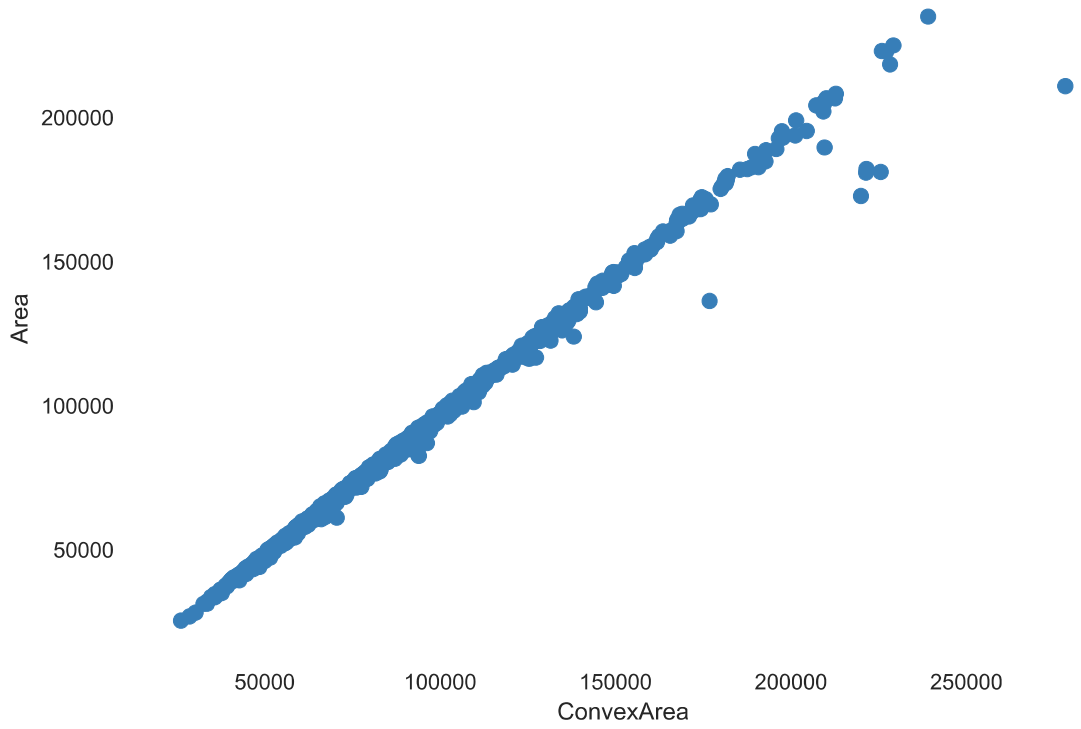

500
450
400
350
300
250
200
150
MinorAxisLength
0.4
0.5
0.6
0.7
0.8
0.9
Eccentricity

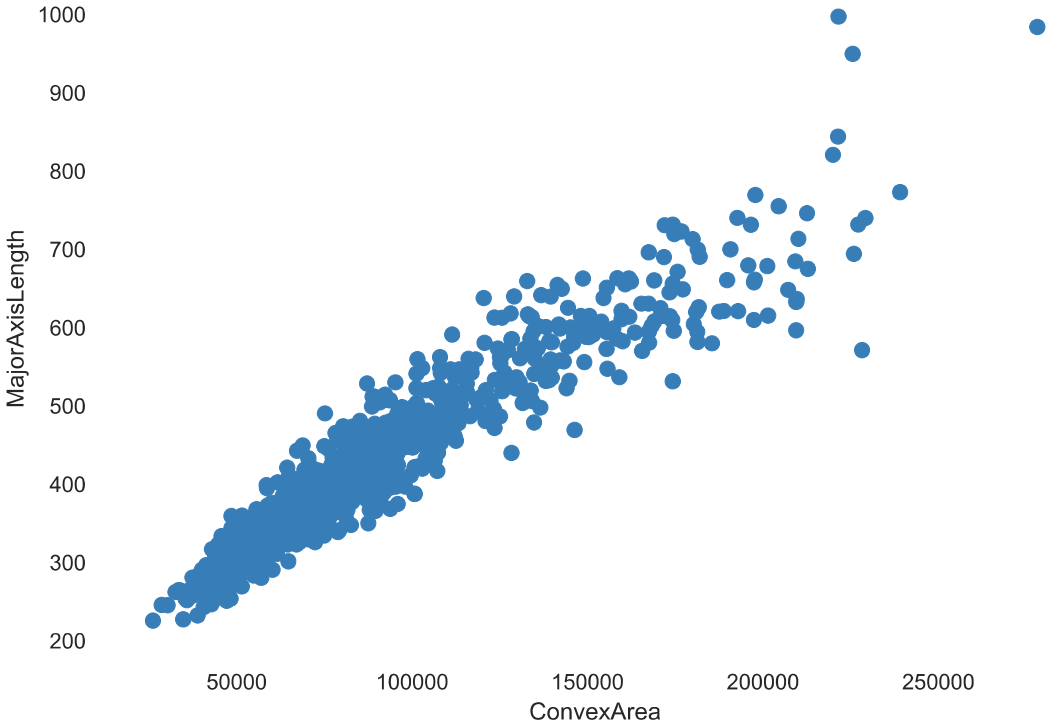

0.9
0.8
0.7
0.6
0.5
0.4
Eccentricity
0.4
0.5
0.6
0.7
0.8
0.9
Eccentricity

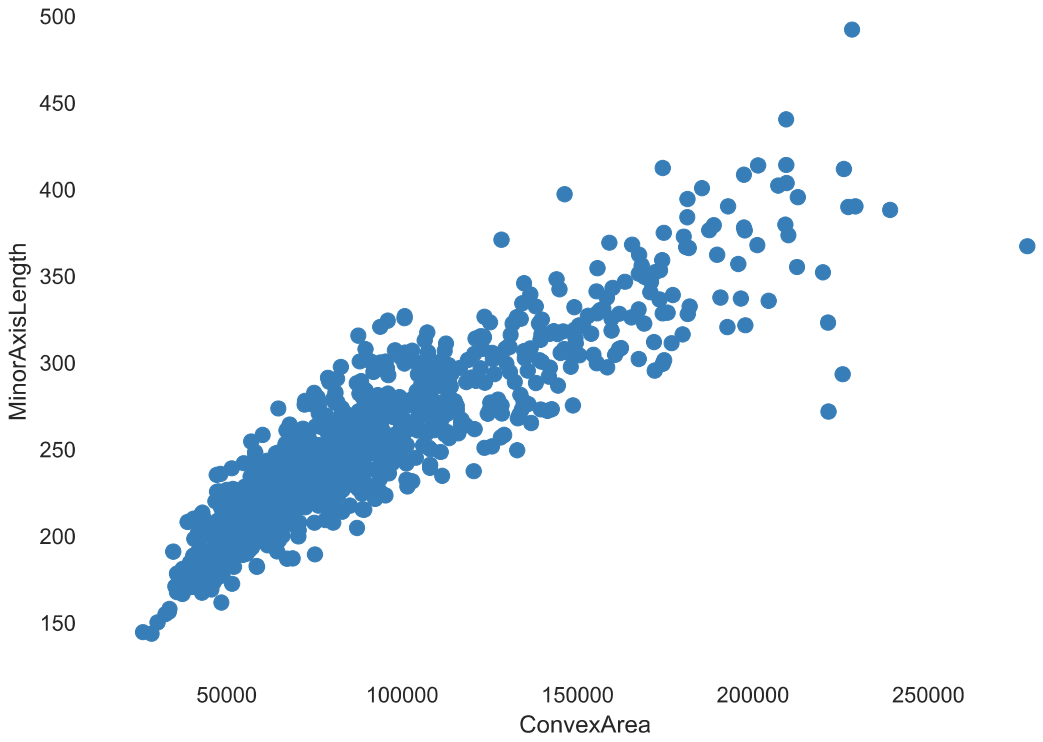

250000
200000
150000
100000
50000
ConvexArea
0.4
0.5
0.6
0.7
0.8
0.9
Eccentricity

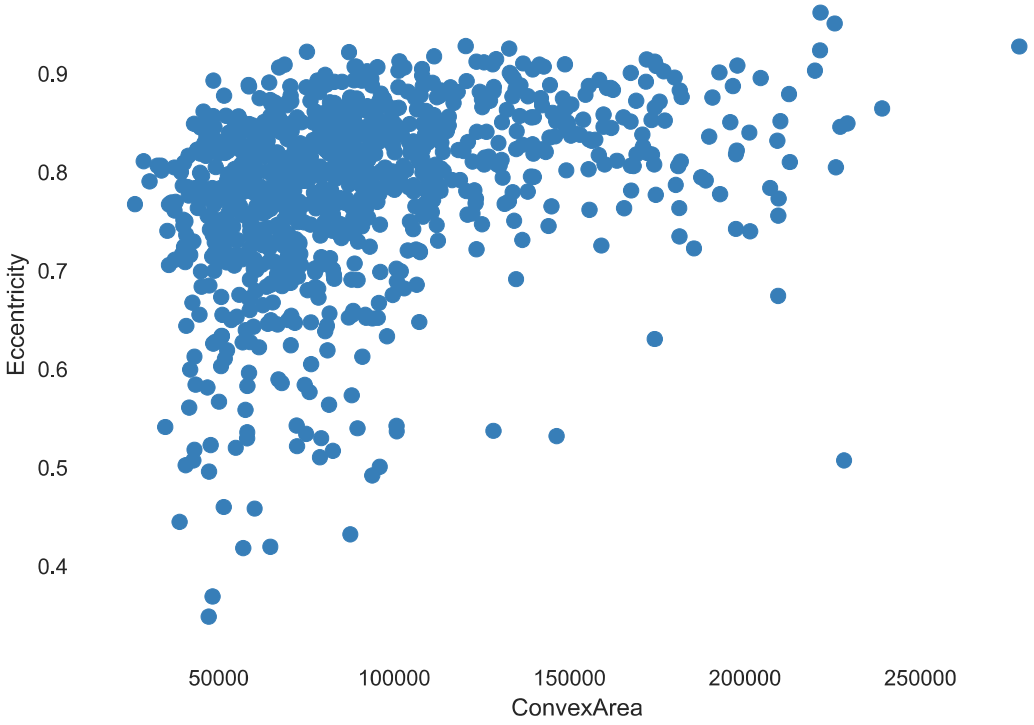

0.8
0.7
0.6
0.5
0.4
Extent
0.4
0.5
0.6
0.7
0.8
0.9
Eccentricity

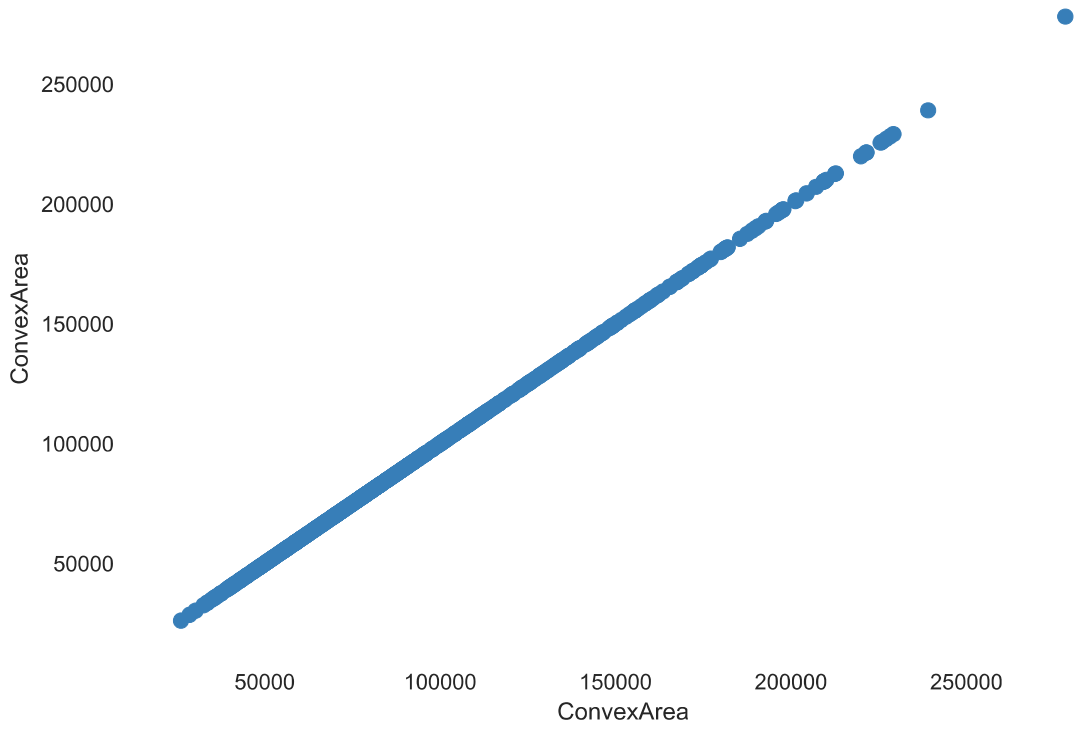

2500
2000
1500
1000
Perimeter
50000
100000
150000
200000
250000
ConvexArea

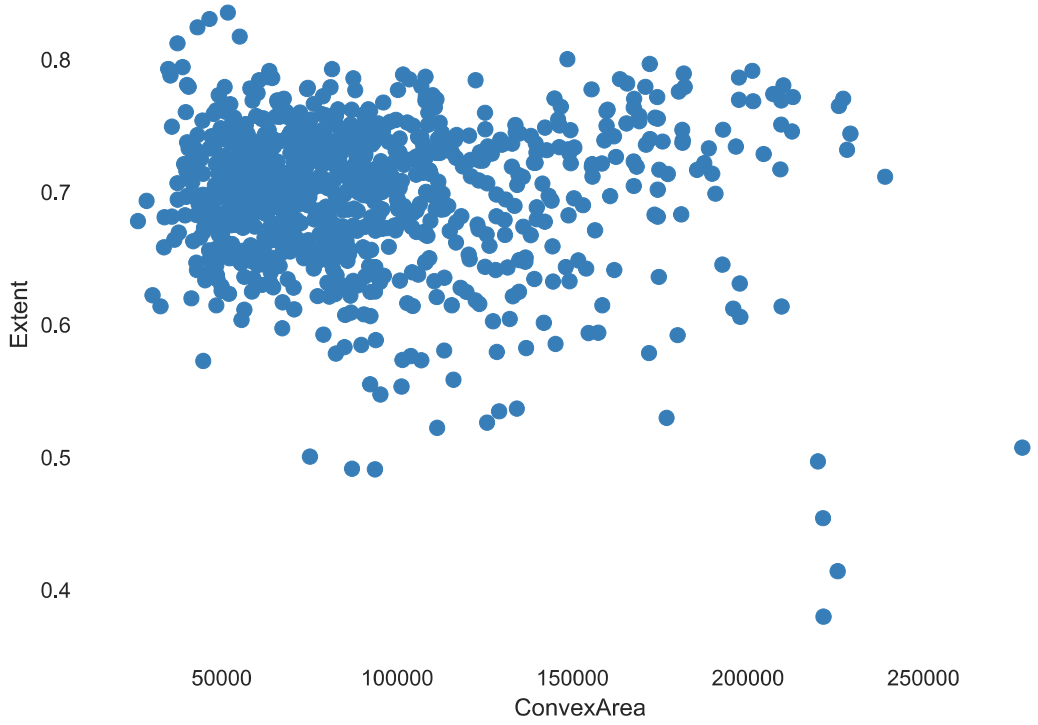

200000
150000
100000
50000
Area
50000
100000
150000
200000
250000
ConvexArea

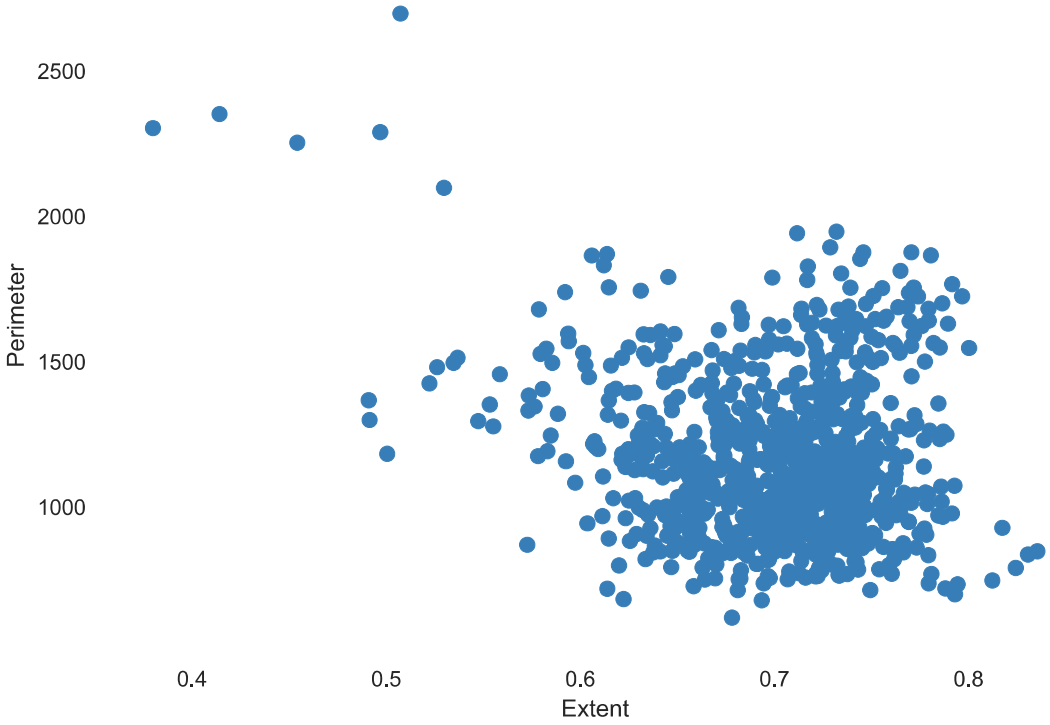

1000
900
800
700
600
500
400
300
200
MajorAxisLength
50000
100000
150000
200000
250000
ConvexArea

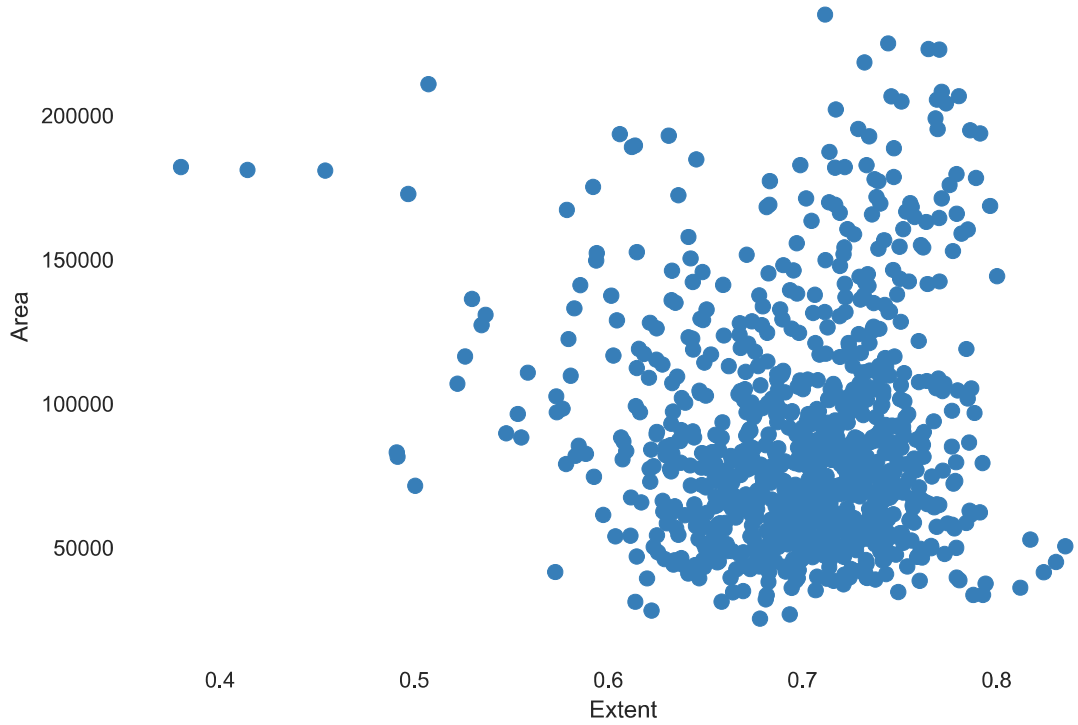

500
450
400
350
300
250
200
150
MinorAxisLength
50000
100000
150000
200000
250000
ConvexArea

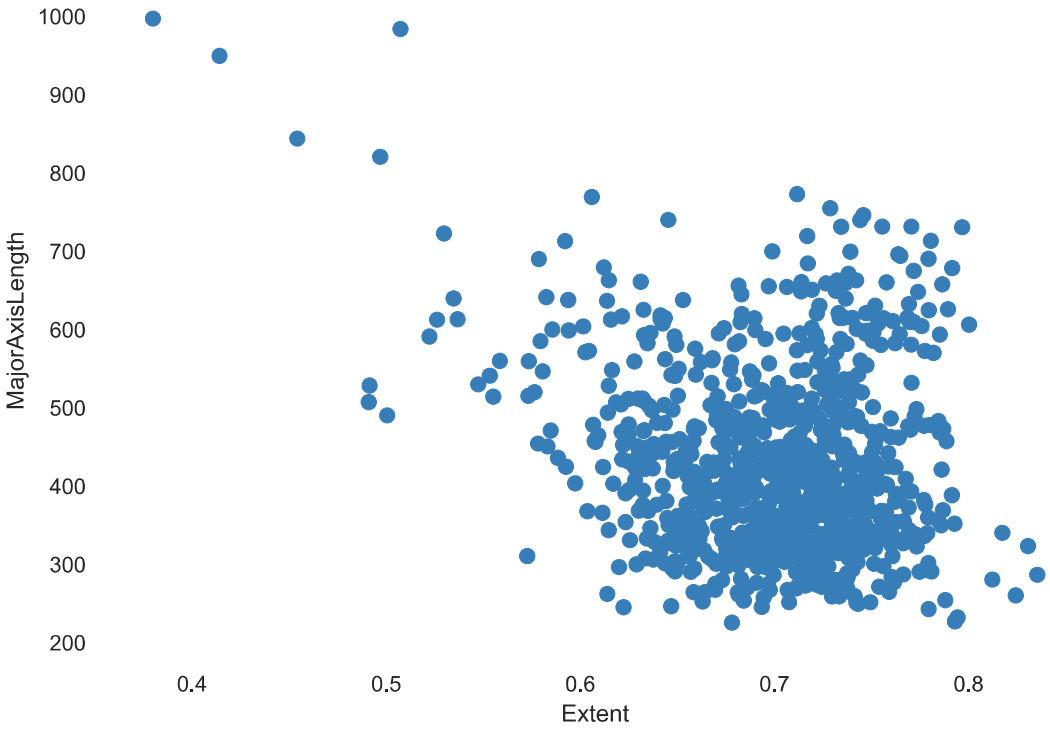

0.9
0.8
0.7
0.6
0.5
0.4
Eccentricity
50000
100000
150000
200000
250000
ConvexArea

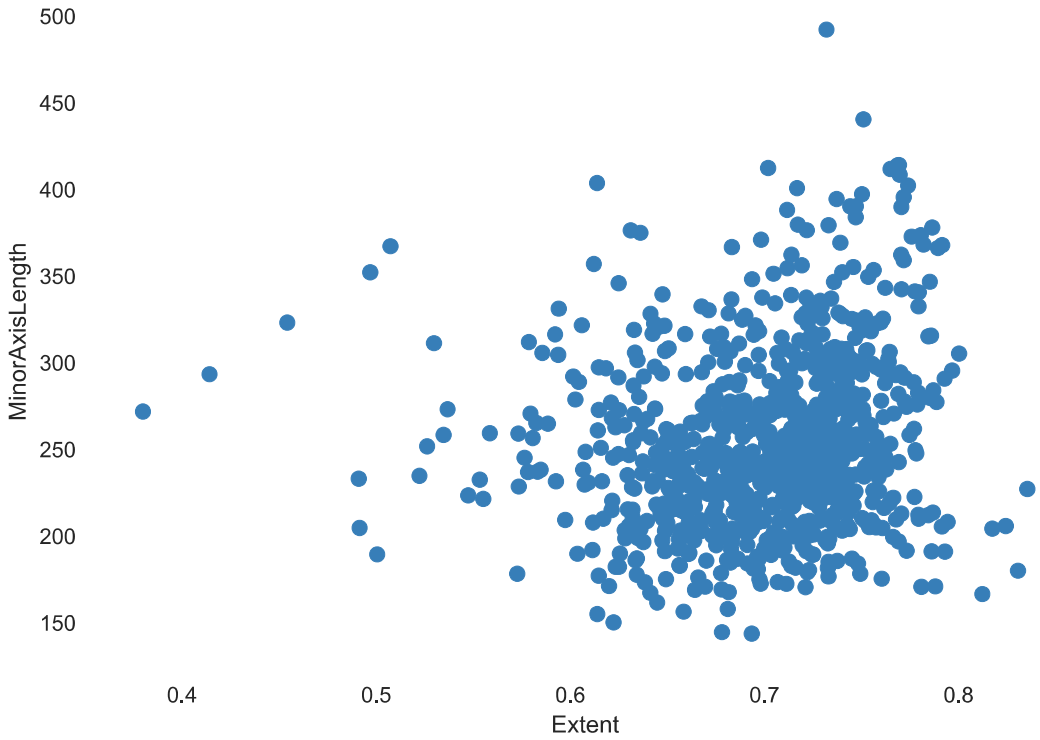

250000
200000
150000
100000
50000
ConvexArea
50000
100000
150000
200000
250000
ConvexArea

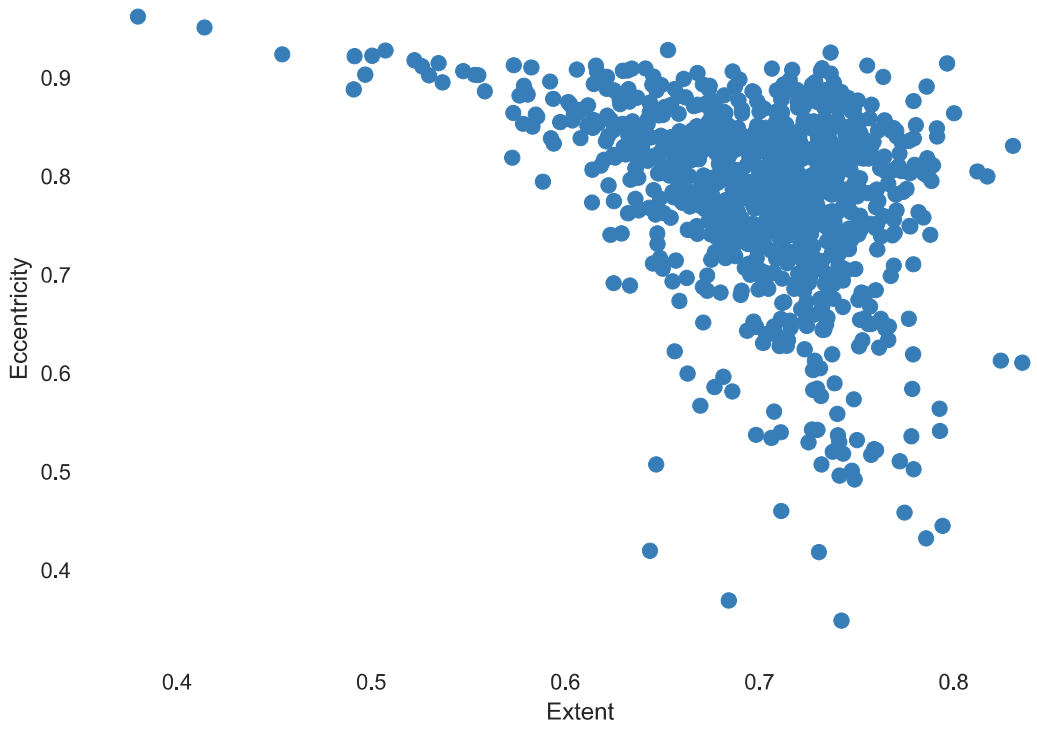

0.8
0.7
0.6
0.5
0.4
Extent
50000
100000
150000
200000
250000
ConvexArea

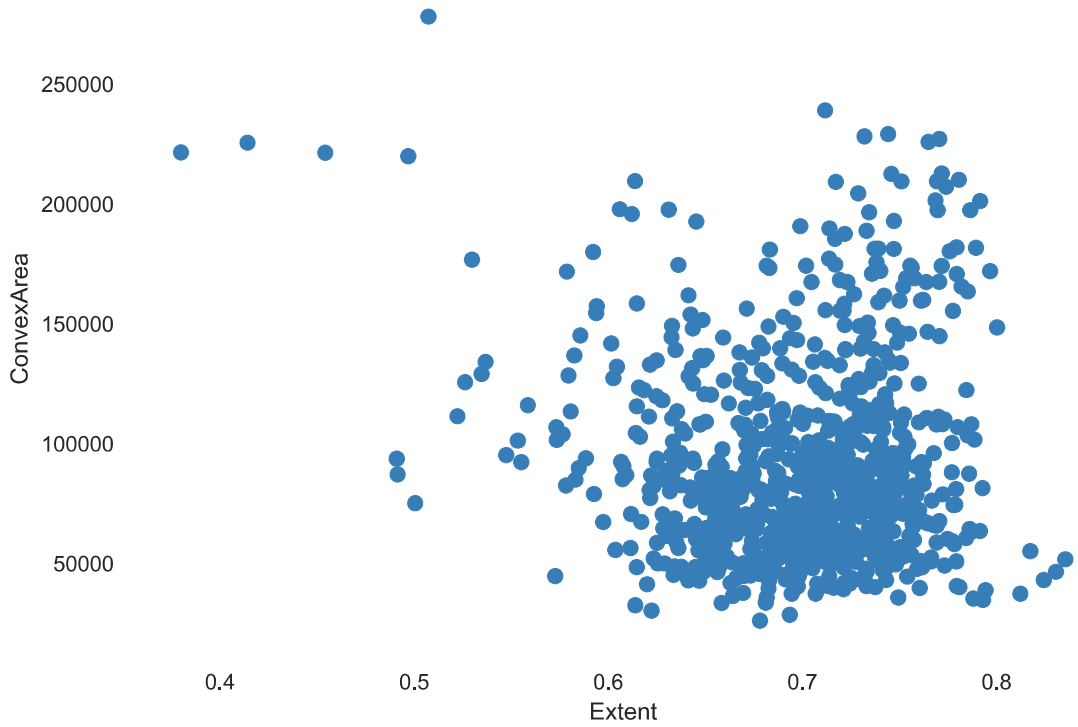

2500
2000
1500
1000
Perimeter
0.4
0.5
0.6
0.7
0.8
Extent

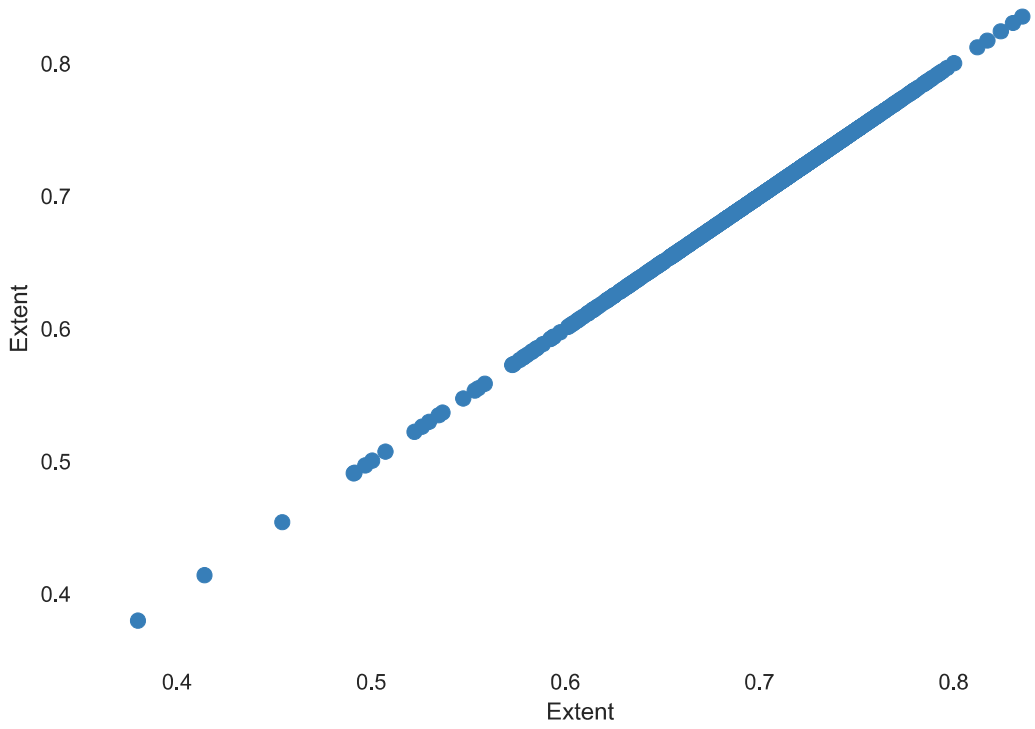

200000
150000
100000
50000
Area
0.4
0.5
0.6
0.7
0.8
Extent

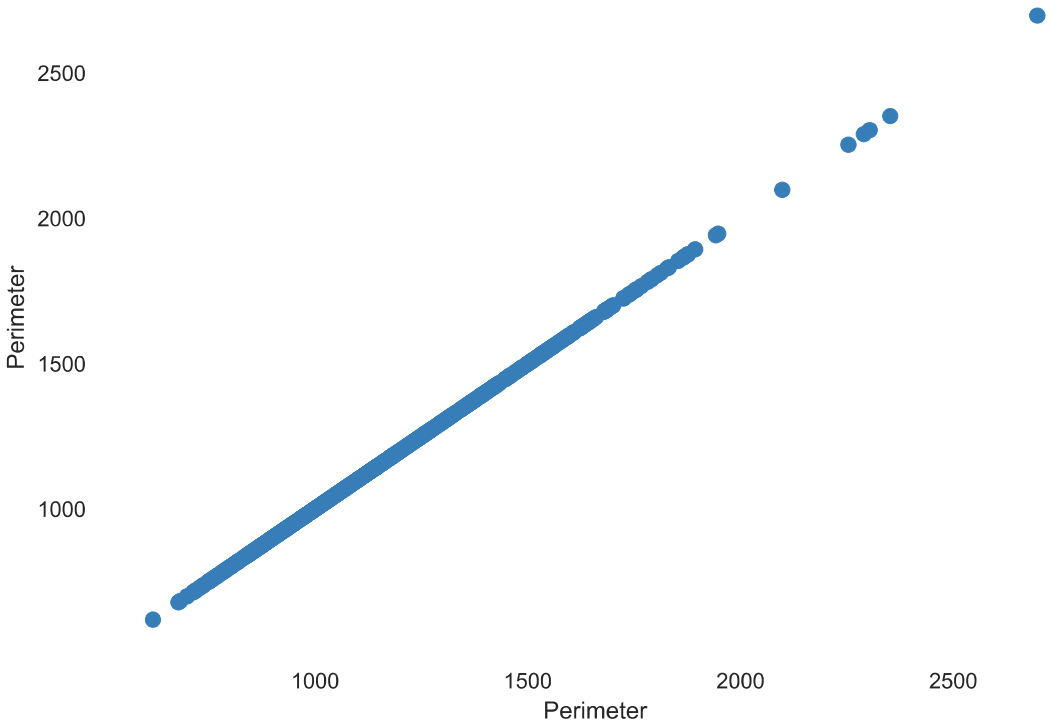

1000
900
800
700
600
500
400
300
200
MajorAxisLength
0.4
0.5
0.6
0.7
0.8
Extent

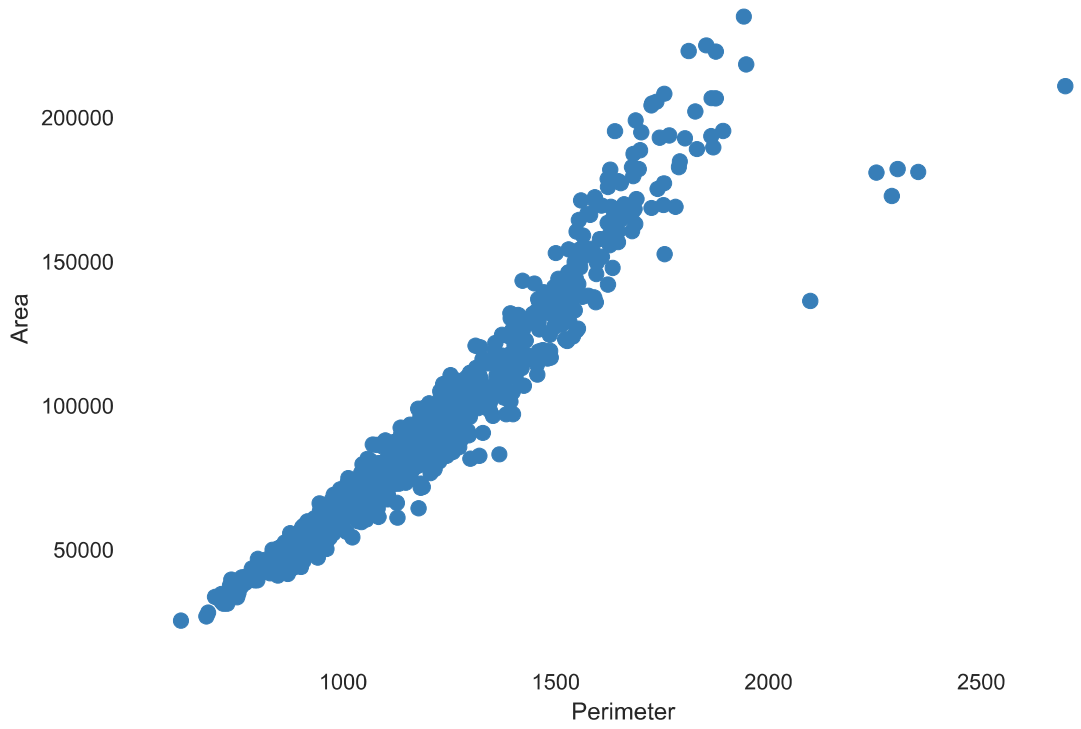

500
450
400
350
300
250
200
150
MinorAxisLength
0.4
0.5
0.6
0.7
0.8
Extent

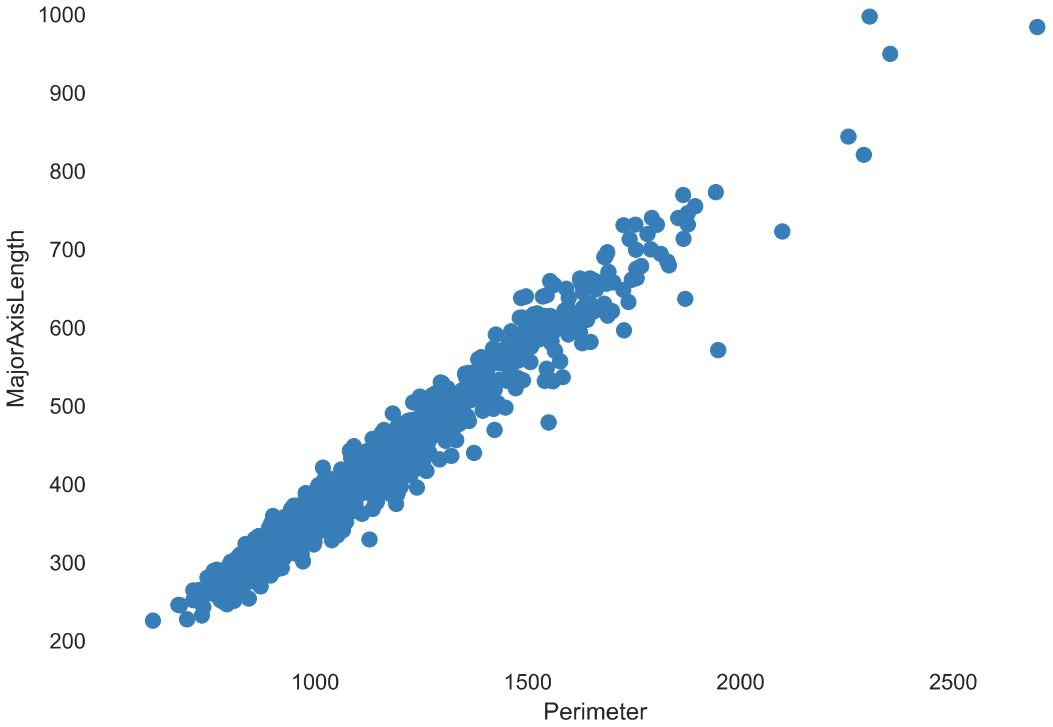

Eccentricity
0.9
0.8
0.7
0.6
0.5
0.4
0.4
0.5
0.6
0.7
0.8
Extent

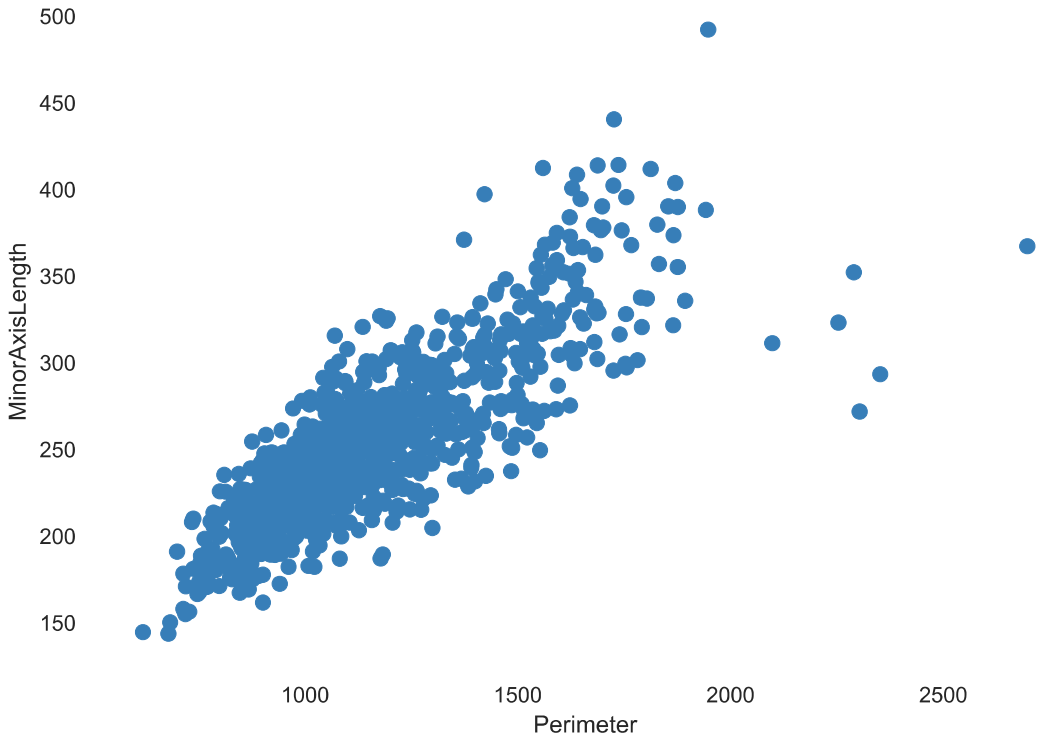

250000
200000
150000
100000
50000
ConvexArea
0.4
0.5
0.6
0.7
0.8
Extent

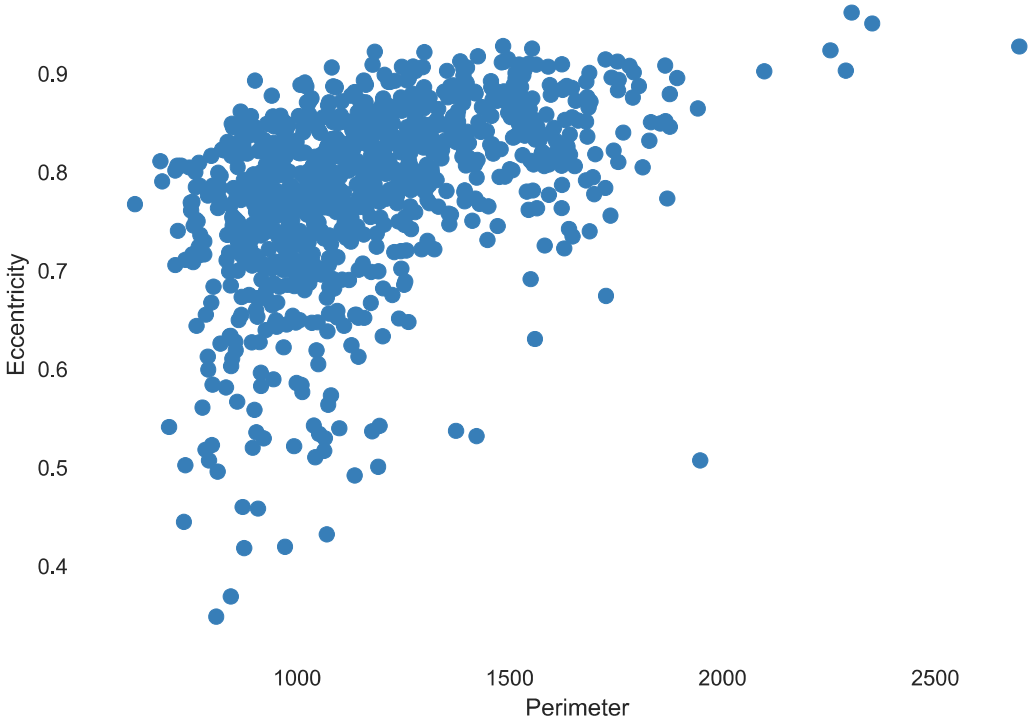

0.8
0.7
0.6
0.5
0.4
Extent
0.4
0.5
0.6
0.7
0.8
Extent

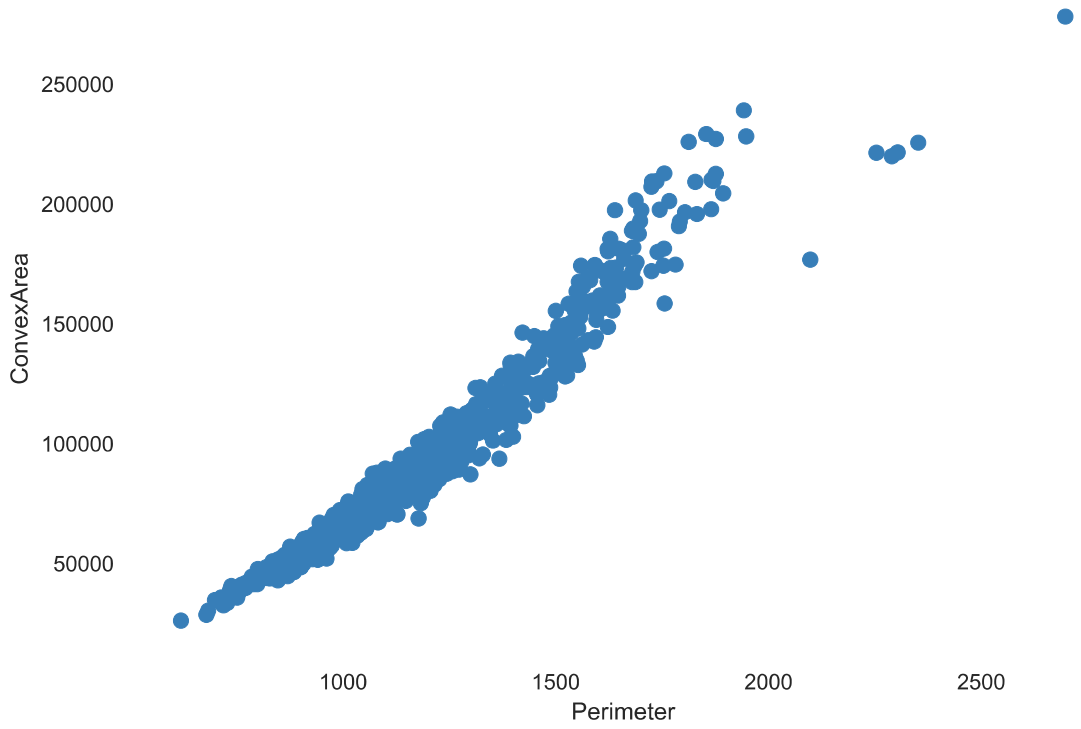

2500
2000
1500
1000
Perimeter
1000
1500
2000
2500
Perimeter

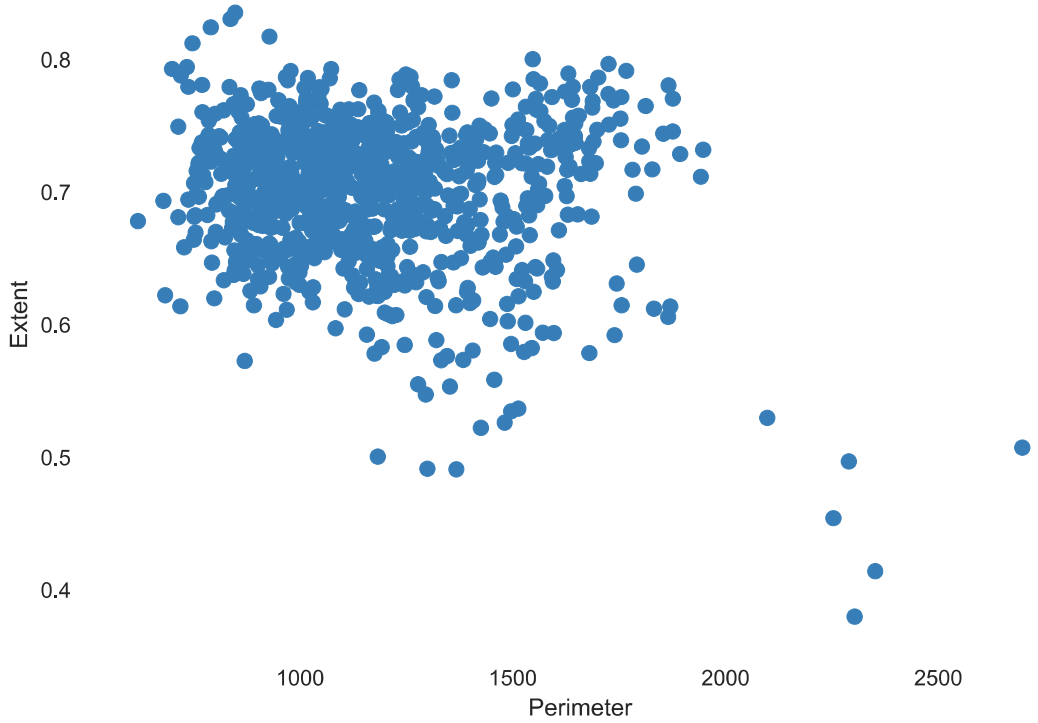

200000
150000
100000
50000
Area
1000
1500
2000
2500
Perimeter


1000
900
800
700
600
500
400
300
200
MajorAxisLength
1000
1500
2000
2500
Perimeter

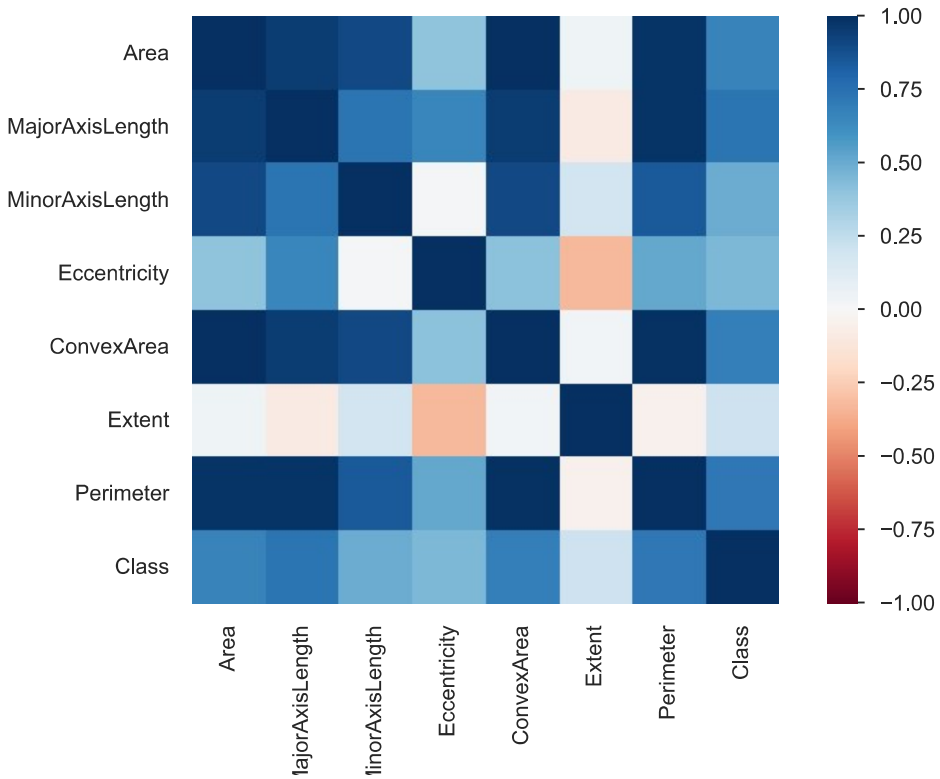

500
450
400
350
300
250
200
150
MinorAxisLength
1000
1500
2000
2500
Perimeter


0.9
0.8
0.7
0.6
0.5
0.4
Eccentricity
1000
1500
2000
2500
Perimeter


250000
200000
150000
100000
50000
ConvexArea
1000
1500
2000
2500
Perimeter

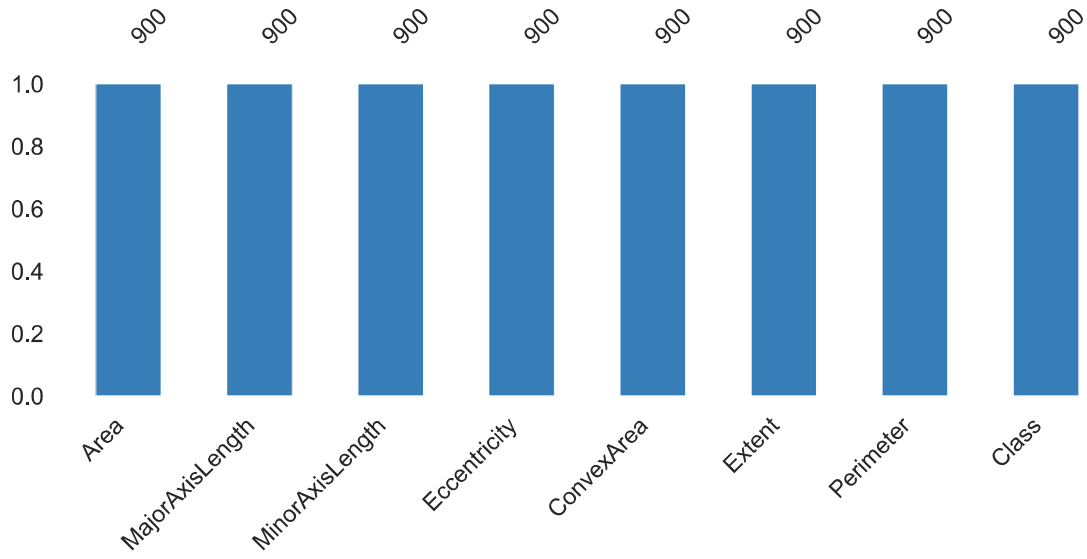

0.8
0.7
0.6
0.5
0.4
Extent
1000
1500
2000
2500
Perimeter

Correlations

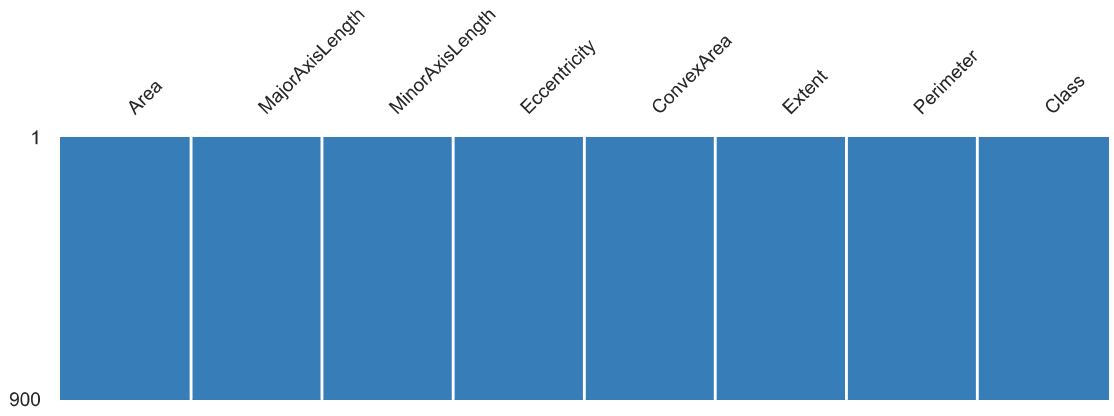


	Area	MajorAxisLength	MinorAxisLength	Eccentricity	ConvexA
Area	1.000	0.949	0.905	0.403	0.999
MajorAxisLength	0.949	1.000	0.733	0.654	0.953
MinorAxisLength	0.905	0.733	1.000	0.012	0.899
Eccentricity	0.403	0.654	0.012	1.000	0.414
ConvexArea	0.999	0.953	0.899	0.414	1.000
Extent	0.043	-0.079	0.192	-0.323	0.028
Perimeter	0.983	0.977	0.837	0.511	0.988
Class	0.671	0.731	0.496	0.452	0.682

Missing values



A simple visualization of nullity by column.



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

Sample

	Area	MajorAxisLength	MinorAxisLength	Eccentricity	ConvexArea	Extent	Per
0	87524	442.246011	253.291155	0.819738	90546	0.758651	118
1	75166	406.690687	243.032436	0.801805	78789	0.684130	112
2	90856	442.267048	266.328318	0.798354	93717	0.637613	120
3	45928	286.540559	208.760042	0.684989	47336	0.699599	844
4	79408	352.190770	290.827533	0.564011	81463	0.792772	107
5	49242	318.125407	200.122120	0.777351	51368	0.658456	881
6	42492	310.146072	176.131449	0.823099	43904	0.665894	823
7	60952	332.455472	235.429835	0.706058	62329	0.743598	933
8	42256	323.189607	172.575926	0.845499	44743	0.698031	849
9	64380	366.964842	227.771615	0.784056	66125	0.664376	981

	Area	MajorAxisLength	MinorAxisLength	Eccentricity	ConvexArea	Extent
890	85646	469.774755	238.539384	0.861490	92673	0.681044
891	107486	462.813134	296.091238	0.768571	108914	0.759967
892	149703	637.873030	304.622532	0.878599	154549	0.593805
893	187391	660.655588	362.315007	0.836205	189799	0.713947
894	115272	511.472036	291.591349	0.821574	119773	0.624760
895	83248	430.077308	247.838695	0.817263	85839	0.668793
896	87350	440.735698	259.293149	0.808629	90899	0.636476
897	99657	431.706981	298.837323	0.721684	106264	0.741099
898	93523	476.344094	254.176054	0.845739	97653	0.658798
899	85609	512.081774	215.271976	0.907345	89197	0.632020

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