

Results

Descriptive Statistics

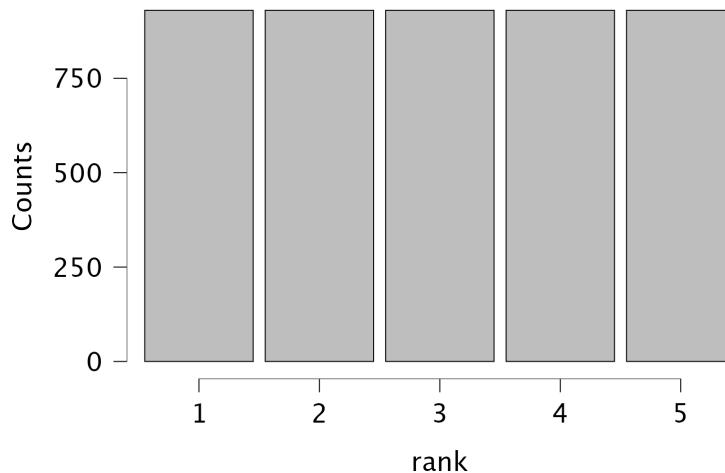
Descriptive Statistics

		Valid	Missing	Mean	Std. Deviation	Shapiro-Wilk	P-value of Shapiro-Wilk	Minimum	Maximum
rank	abalone	4650	0	3.000	1.414	0.888	< .001	1.000	5.000
rank	adult	4650	0	2.935	1.436	0.883	< .001	1.000	5.000
rank	air_quality	4650	0	3.000	1.414	0.888	< .001	1.000	5.000
rank	bike	4650	0	3.000	1.414	0.888	< .001	1.000	5.000
rank	car	4650	0	3.000	1.414	0.888	< .001	1.000	5.000
rank	fish_toxicity	4650	0	3.000	1.414	0.888	< .001	1.000	5.000
rank	forest_fires	4650	0	3.000	1.414	0.888	< .001	1.000	5.000
rank	housing	4650	0	3.000	1.414	0.888	< .001	1.000	5.000
rank	iris	4650	0	3.000	1.414	0.888	< .001	1.000	5.000
rank	mushroom	4650	0	2.994	1.420	0.887	< .001	1.000	5.000
rank	parkinsons	4650	0	3.000	1.414	0.888	< .001	1.000	5.000
rank	student_performance	4650	0	3.000	1.414	0.888	< .001	1.000	5.000
rank	wine_quality	4650	0	3.000	1.414	0.888	< .001	1.000	5.000
rank	bank	4650	0	3.000	1.414	0.888	< .001	1.000	5.000
rank	diabetic	4650	0	3.000	1.414	0.888	< .001	1.000	5.000
test_loss	abalone	4650	0	2.360	0.616	0.567	< .001	1.938	13.960
test_loss	adult	4650	0	609.398	4695.240	0.088	< .001	0.314	150251.230
test_loss	air_quality	4650	0	0.281	0.050	0.435	< .001	0.242	0.746
test_loss	bike	4650	0	0.113	0.073	0.651	< .001	0.047	0.662
test_loss	car	4650	0	0.279	0.303	0.417	< .001	0.077	2.819
test_loss	fish_toxicity	4650	0	0.113	0.043	0.369	< .001	0.080	0.538
test_loss	forest_fires	4650	0	0.093	0.105	0.622	< .001	0.012	0.885
test_loss	housing	4650	0	0.113	0.053	0.572	< .001	0.060	0.583
test_loss	iris	4650	0	0.300	0.579	0.349	< .001	0.001	19.811
test_loss	mushroom	4650	0	0.372	3.509	0.071	< .001	0.000	135.379
test_loss	parkinsons	4650	0	0.074	0.056	0.275	< .001	0.054	0.659
test_loss	student_performance	4650	0	0.277	0.127	0.776	< .001	0.145	0.614
test_loss	wine_quality	4650	0	1.149	0.250	0.329	< .001	1.009	2.996
test_loss	bank	4650	0	0.266	0.210	0.177	< .001	0.201	9.872
test_loss	diabetic	4650	0	1.929	7.857	0.090	< .001	0.884	371.826

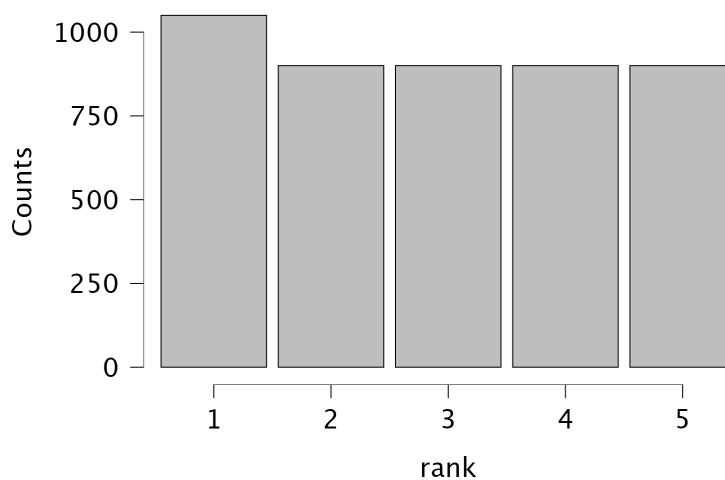
Distribution Plots

rank

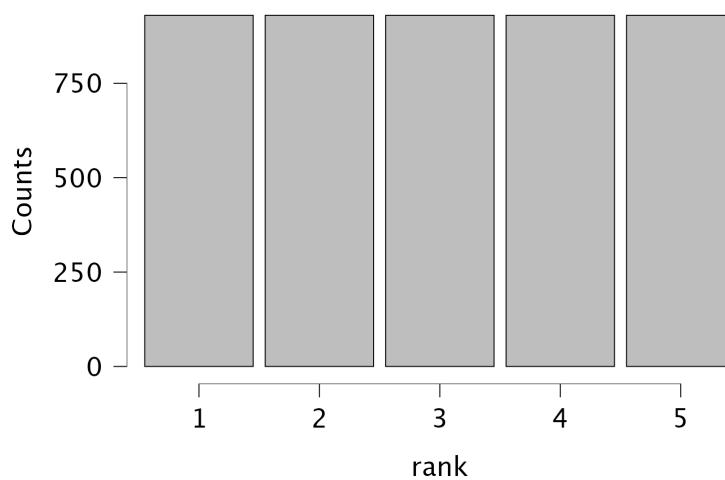
abalone



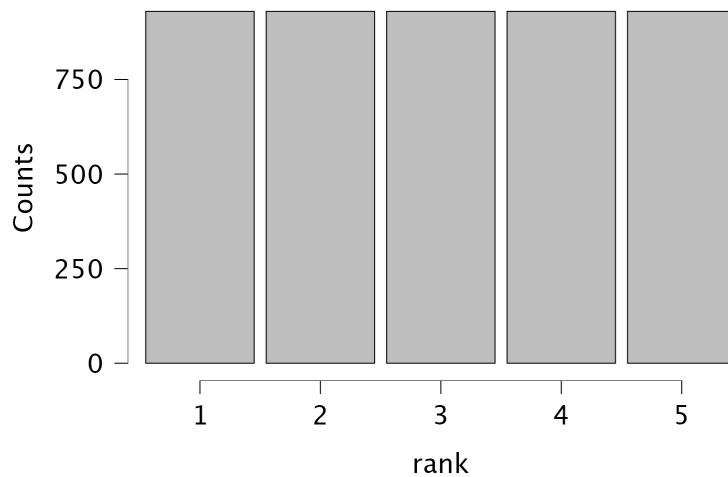
adult



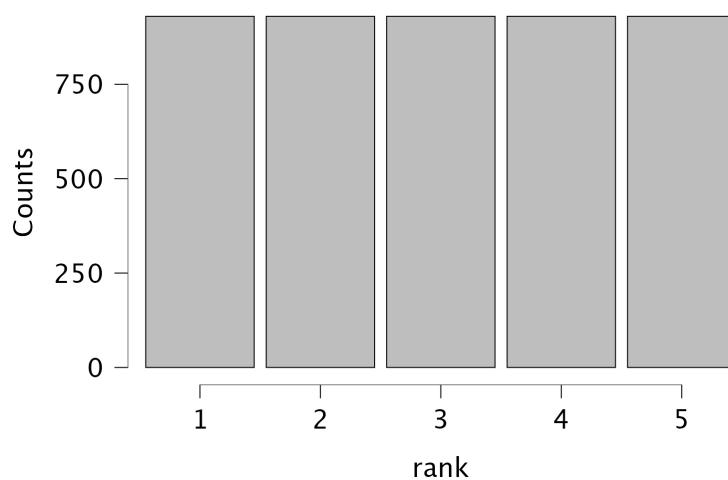
air_quality



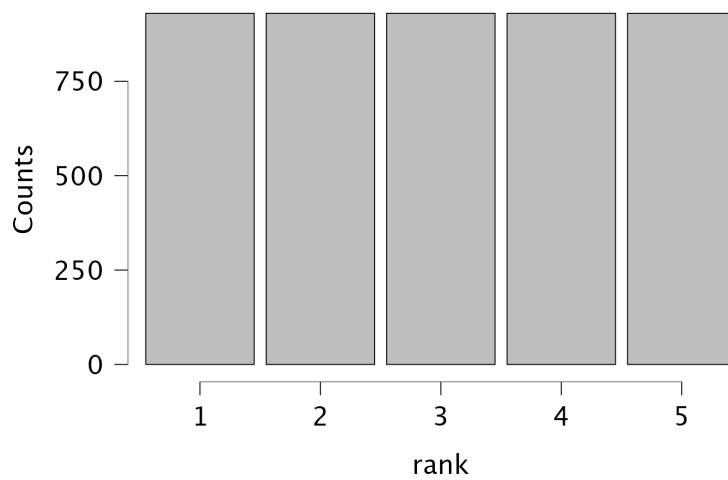
bike



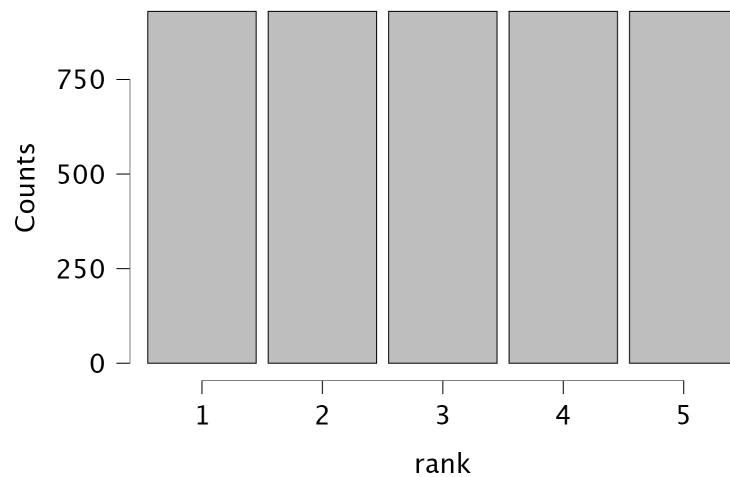
car



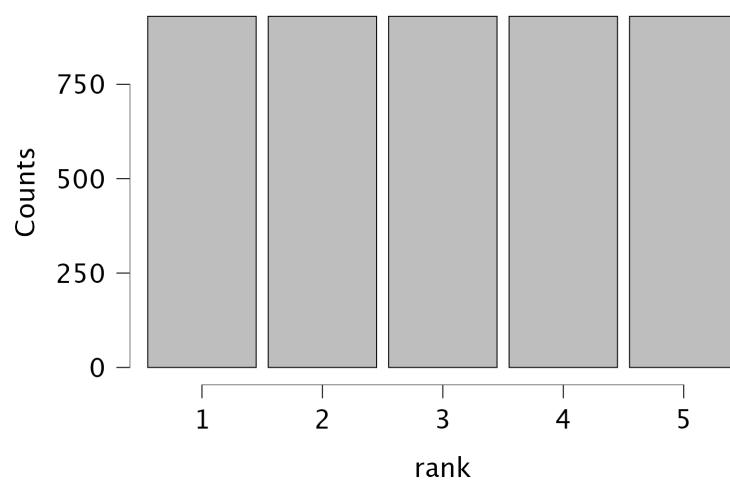
fish_toxicity



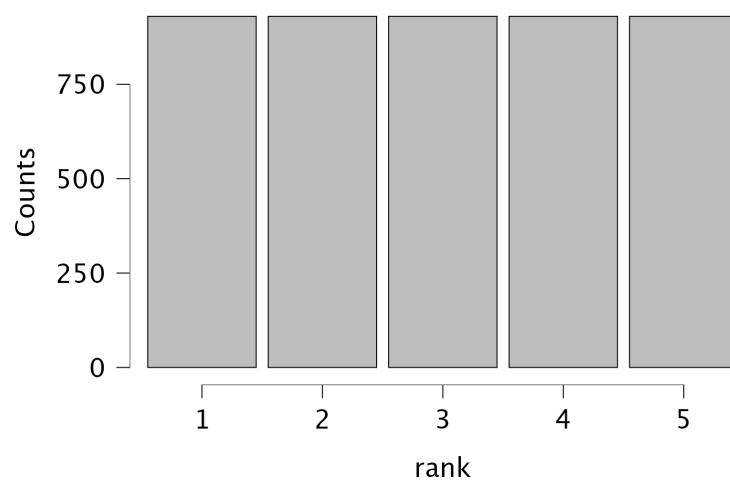
forest_fires



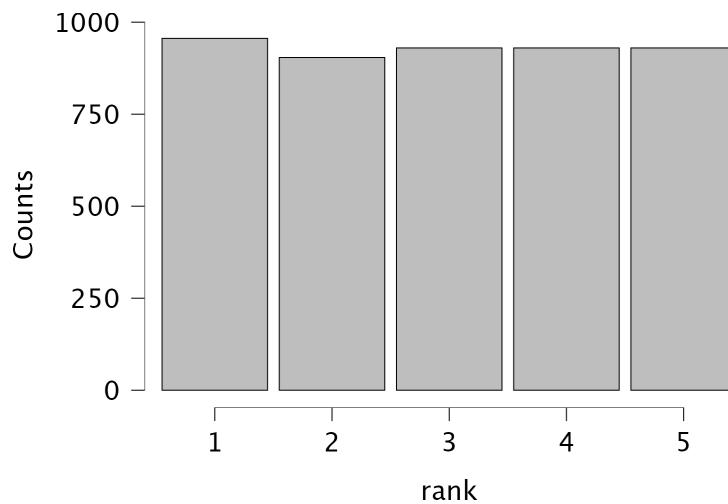
housing



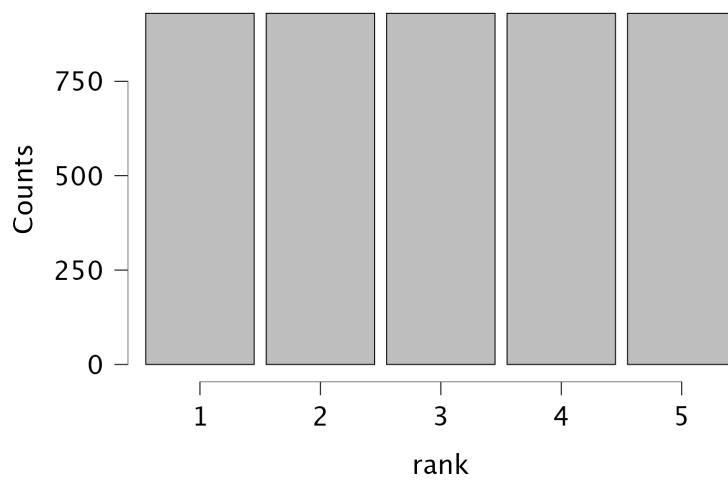
iris



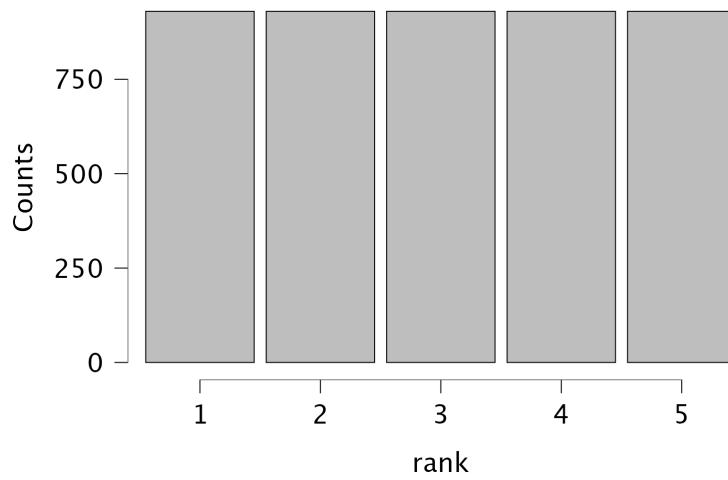
mushroom



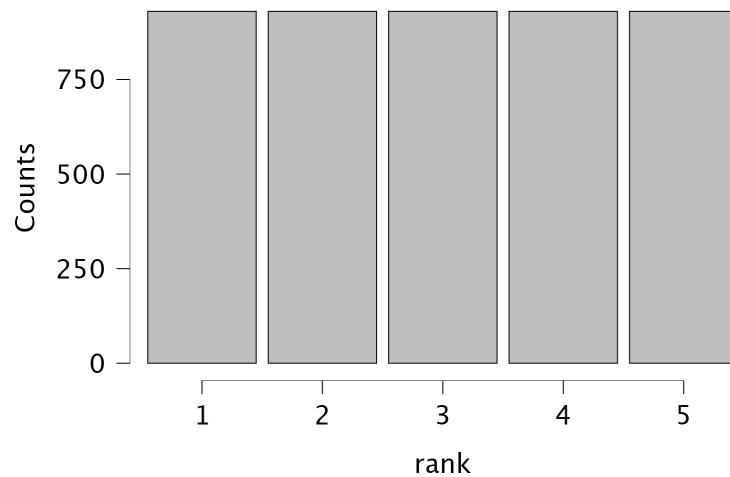
parkinsons



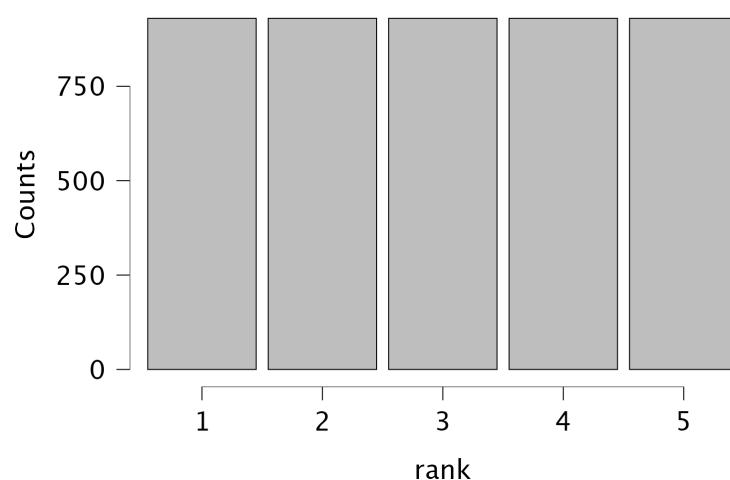
student_performance



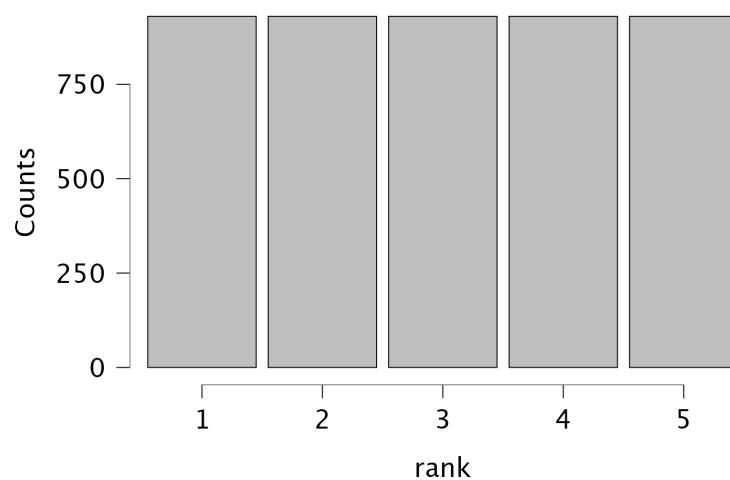
wine_quality



bank

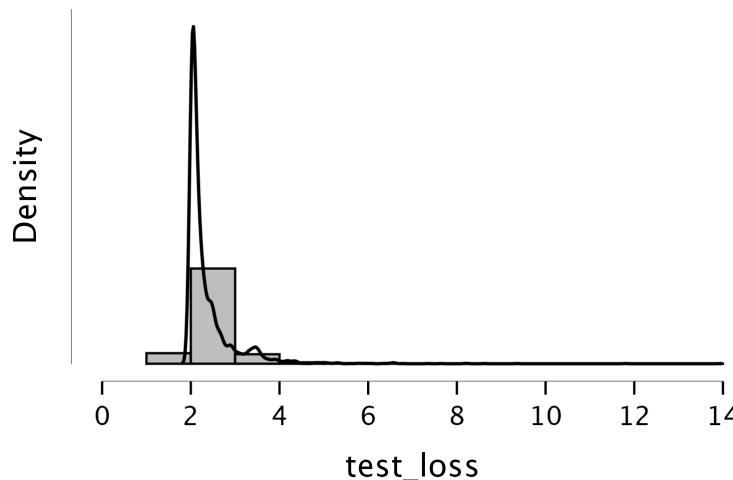


diabetic

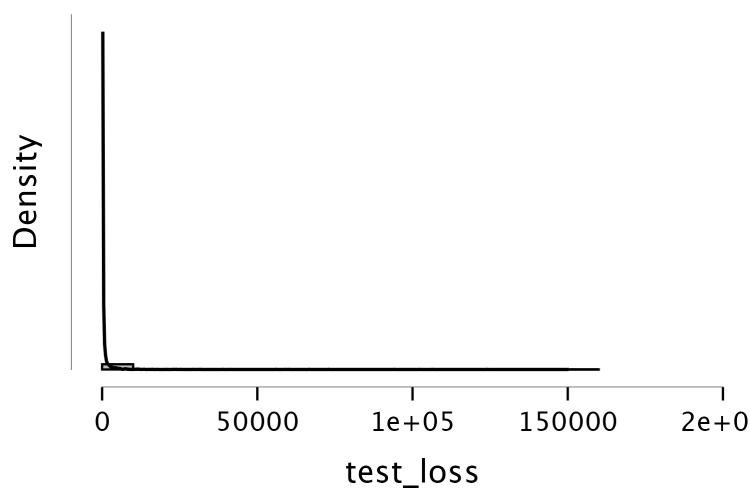


test_loss

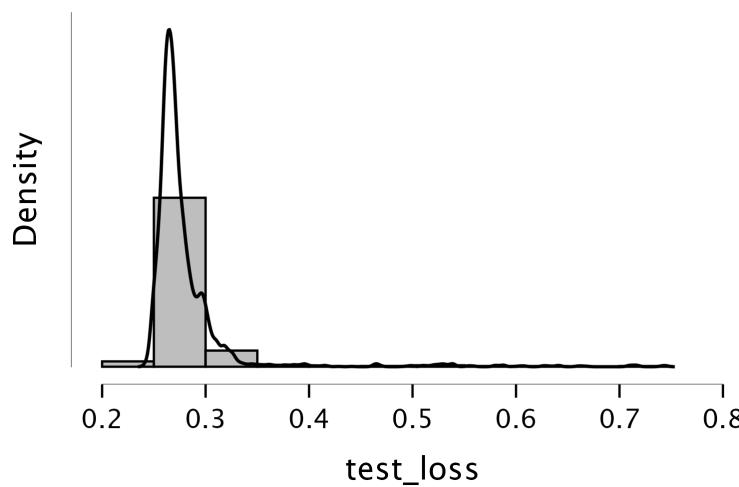
abalone



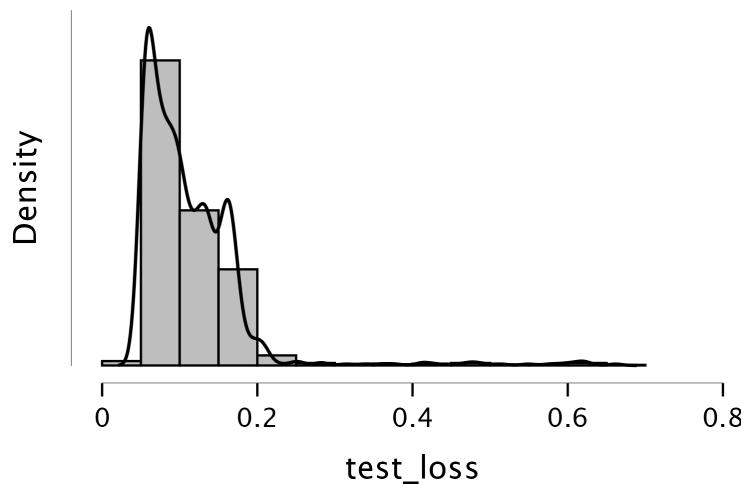
adult



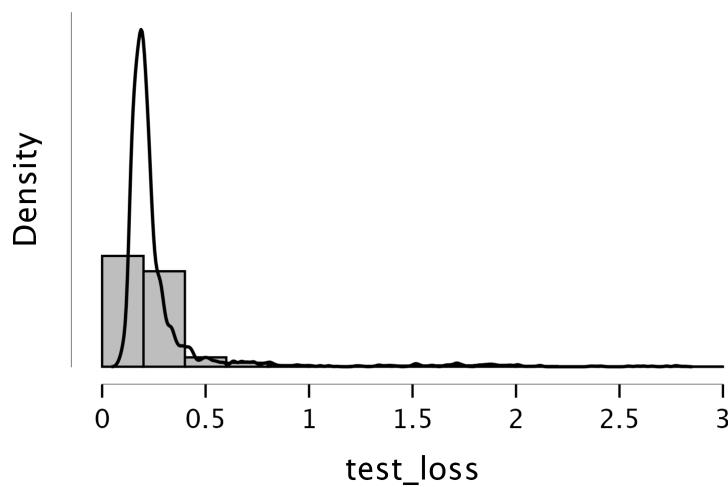
air_quality



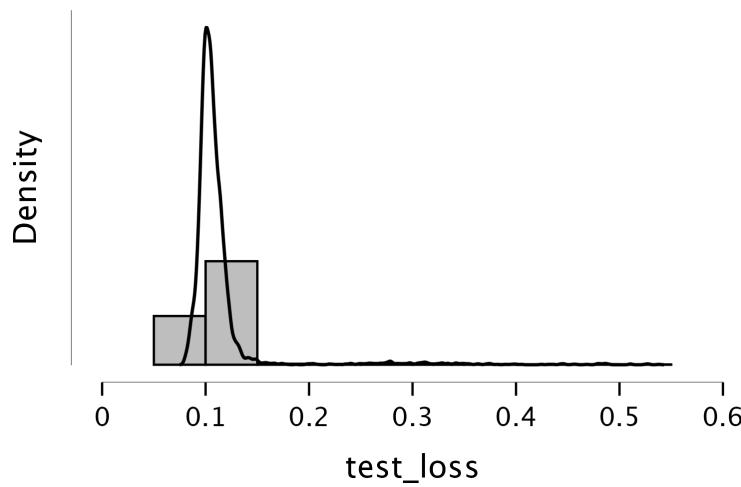
bike



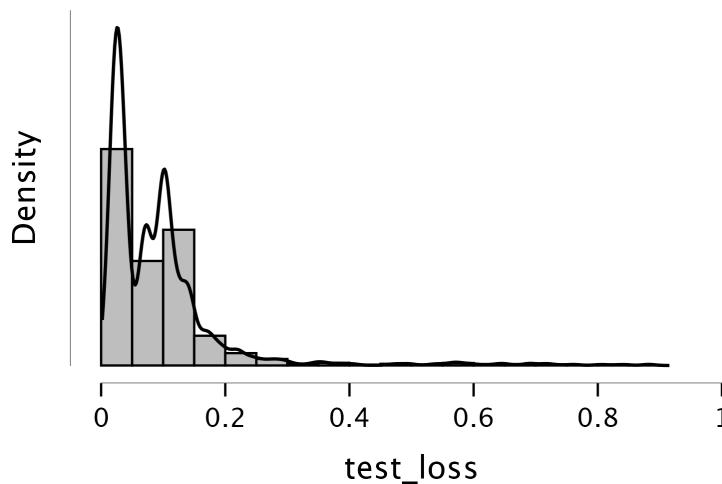
car



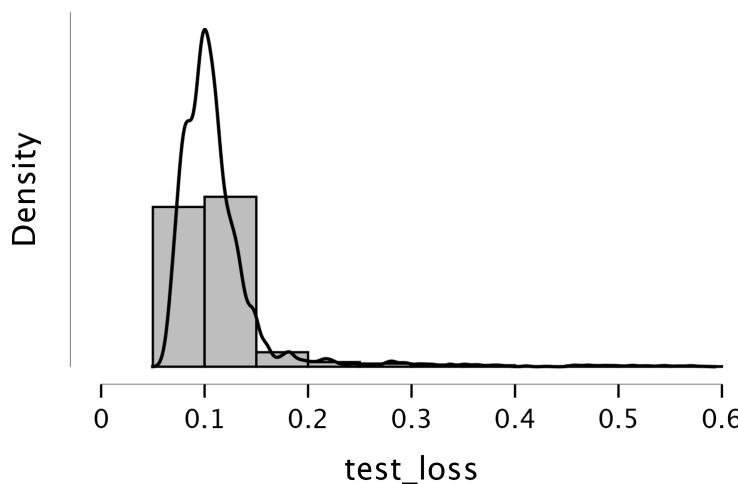
fish_toxicity



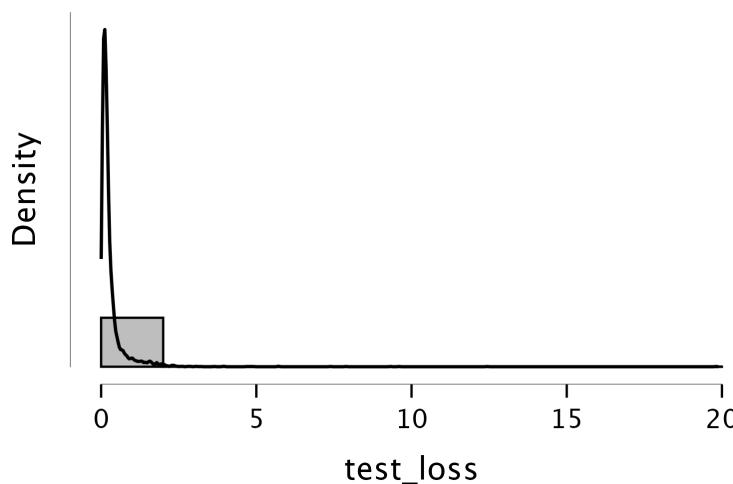
`forest_fires`



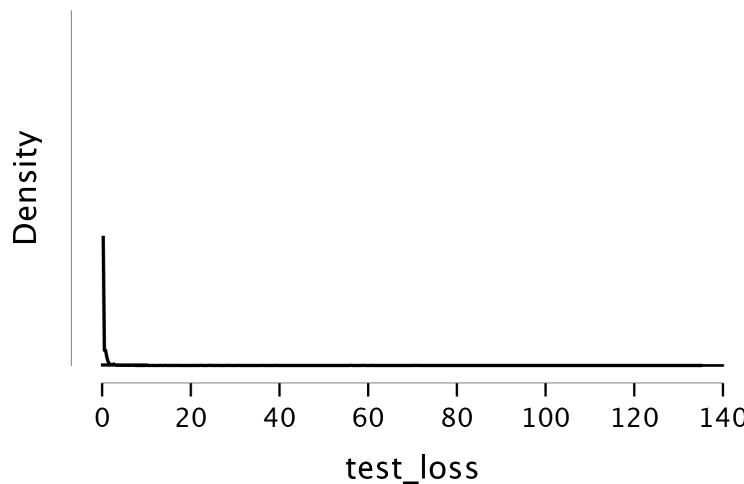
`housing`



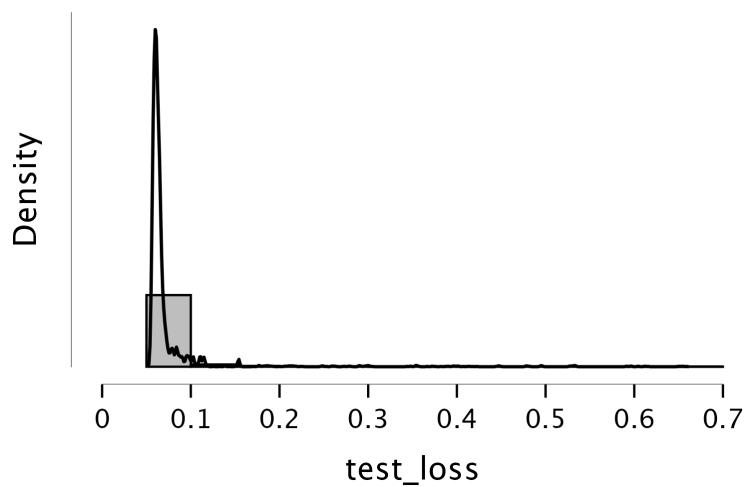
`iris`



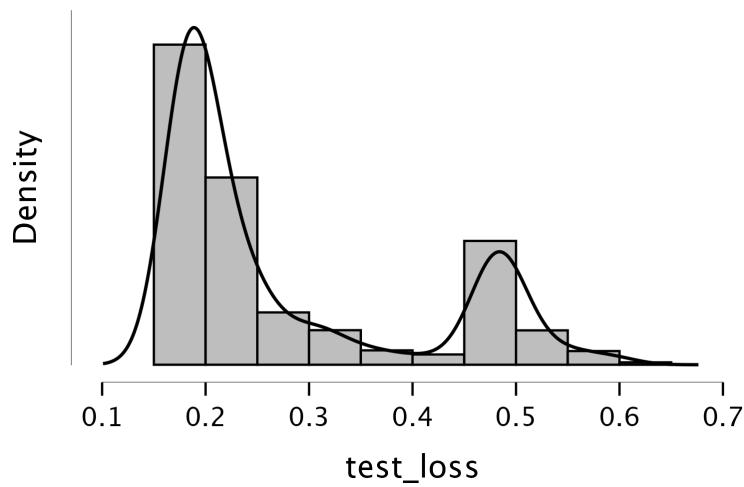
mushroom



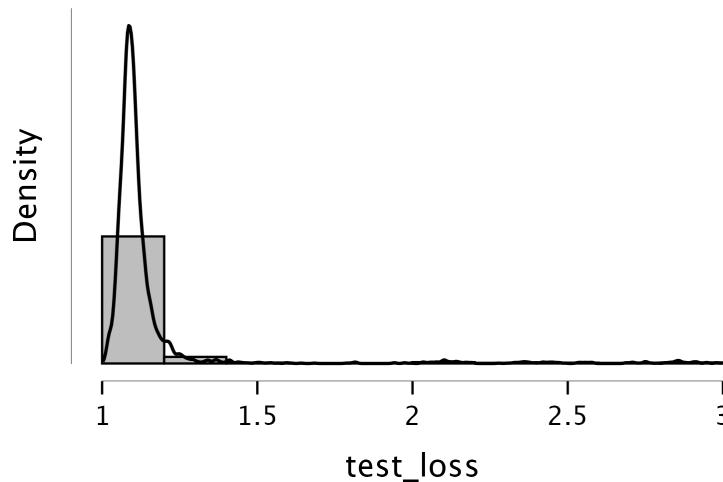
parkinsons



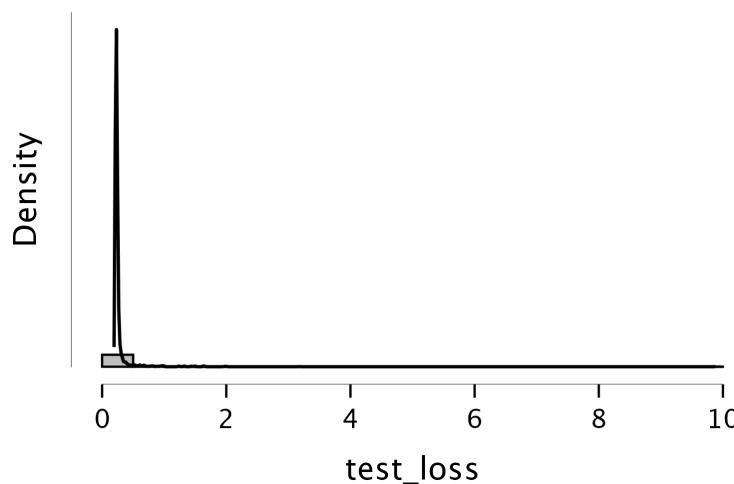
student_performance



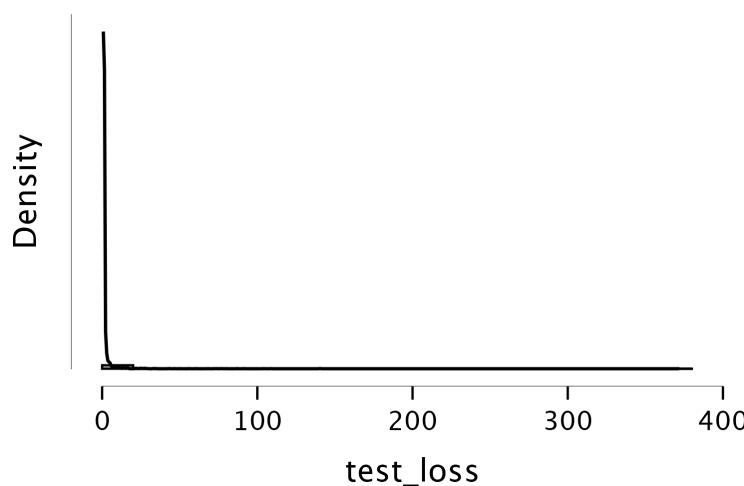
wine_quality



bank



diabetic



ANOVA

ANOVA – rank

Cases	Sum of Squares	df	Mean Square	F	p
dataset	17.977	14	1.284	0.696	0.781
burn_in	6332.995	4	1583.249	858.000	< .001
dataset * burn_in	4955.754	56	88.496	47.958	< .001
Residuals	128569.751	69675	1.845		

Note. Type III Sum of Squares

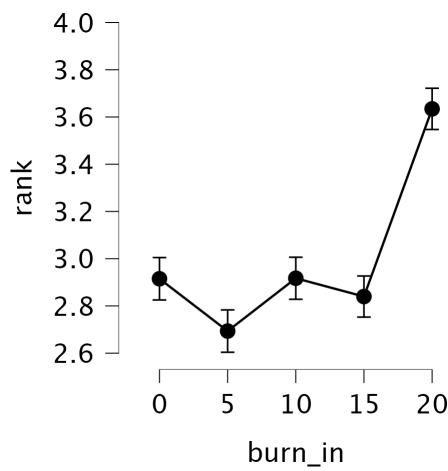
Descriptives

Descriptives – rank

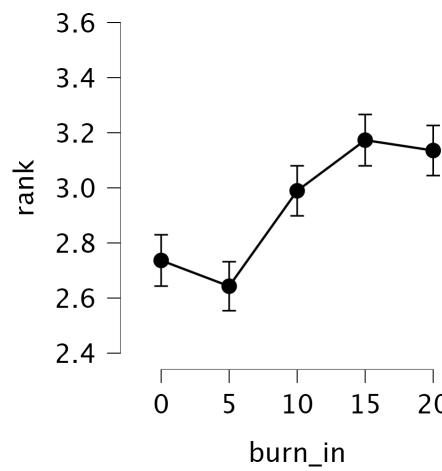
dataset	burn_in	Mean	SD	N
abalone	0	2.915	1.395	930
	10	2.917	1.384	930
	15	2.840	1.352	930
	20	3.634	1.357	930
	5	2.694	1.393	930
	0	2.737	1.447	930
adult	10	2.989	1.411	930
	15	3.173	1.448	930
	20	3.135	1.413	930
	5	2.643	1.382	930
	0	2.910	1.441	930
	10	2.770	1.344	930
air_quality	15	3.372	1.314	930
	20	3.267	1.412	930
	5	2.682	1.427	930
	0	2.730	1.321	930
	10	2.857	1.485	930
	15	2.931	1.384	930
bank	20	3.662	1.339	930
	5	2.819	1.335	930
	0	1.924	0.981	930
	10	2.924	1.274	930
	15	3.505	1.186	930
	20	4.467	0.851	930
bike	5	2.181	1.024	930
	0	2.362	1.231	930
	10	3.298	1.392	930
	15	3.380	1.271	930
	20	3.372	1.401	930
	5	2.588	1.425	930
diabetic	0	2.405	1.358	930
	10	3.240	1.372	930
	15	3.232	1.326	930
	20	3.422	1.396	930
	5	2.701	1.358	930
	0	3.063	1.459	930
fish_toxicity	10	2.820	1.513	930
	15	3.102	1.346	930
	20	3.220	1.295	930
	5	2.794	1.402	930
	0	2.816	1.330	930
	10	2.865	1.409	930
forest_fires	15	3.377	1.398	930
	20	3.092	1.471	930
	5	2.849	1.383	930
	0	2.690	1.384	930
	10	2.997	1.442	930
	15	3.058	1.344	930
housing	20	3.303	1.510	930
	5	2.952	1.318	930
	0	2.947	1.461	930
	10	3.190	1.449	930
	15	2.771	1.369	930
	20	3.119	1.387	930
iris	5	2.972	1.367	930
	0	2.206	1.249	930
	10	3.161	1.297	930
	15	3.144	1.401	930
	20	3.523	1.501	930
	5	2.938	1.298	930
mushroom	0	2.280	1.282	930
	10	2.706	1.314	930
	15	3.397	1.417	930
	20	3.728	1.330	930
	5	2.889	1.247	930
	0	2.145	1.205	930
parkinsons	10	3.097	1.352	930
	15	3.458	1.354	930
	20	3.605	1.231	930
	5	2.695	1.403	930
	0	2.970	1.437	930
	10	2.663	1.419	930
student_performance	15	3.125	1.398	930
	20	3.231	1.447	930
	5	3.011	1.304	930

Descriptives plots

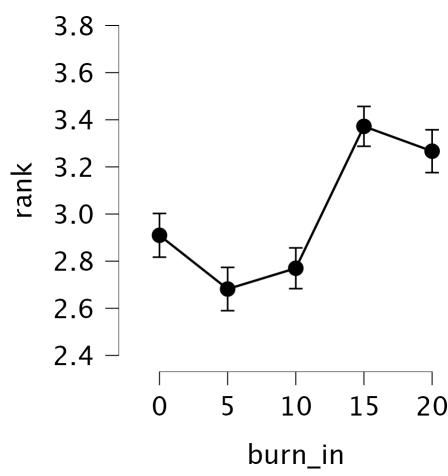
dataset: abalone



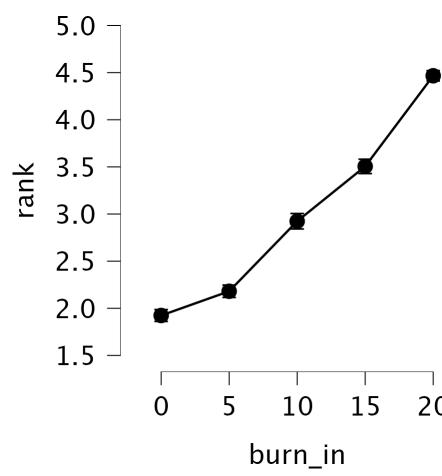
dataset: adult



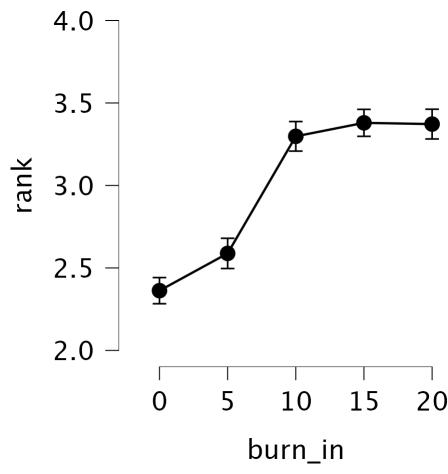
dataset: air_quality



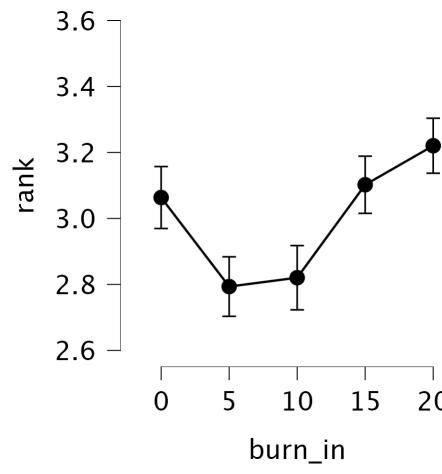
dataset: bike



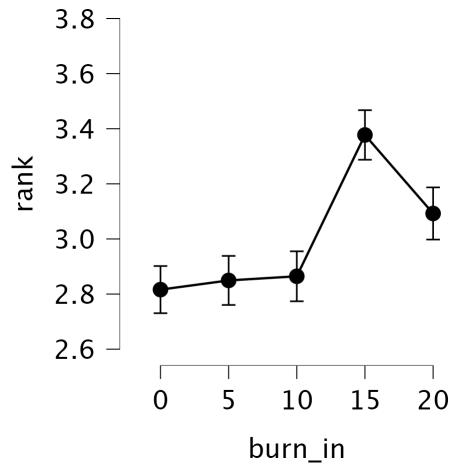
dataset: car



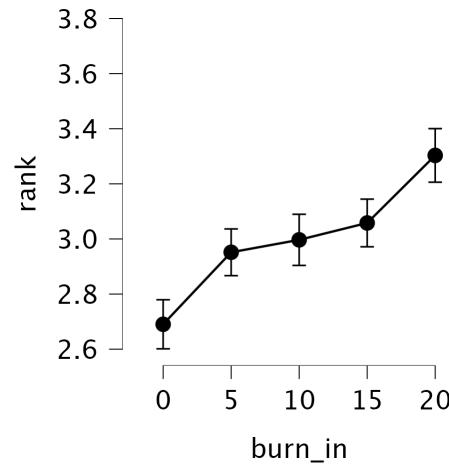
dataset: fish_toxicity



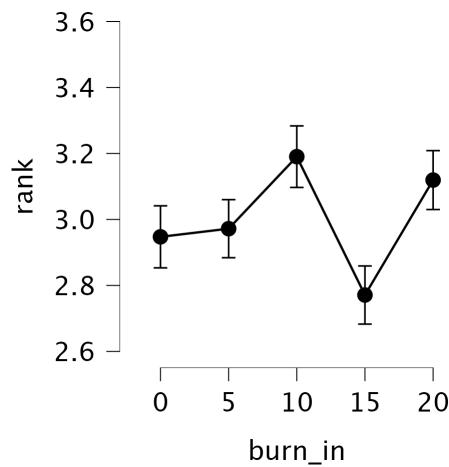
dataset: forest_fires



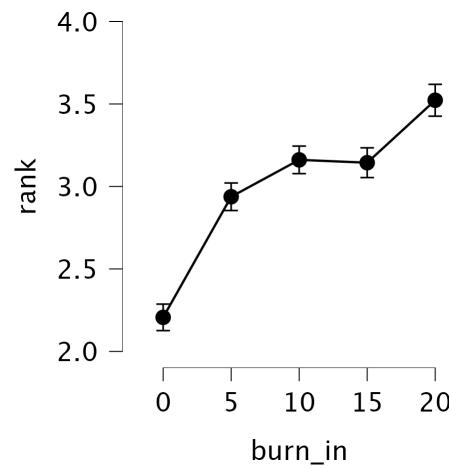
dataset: housing



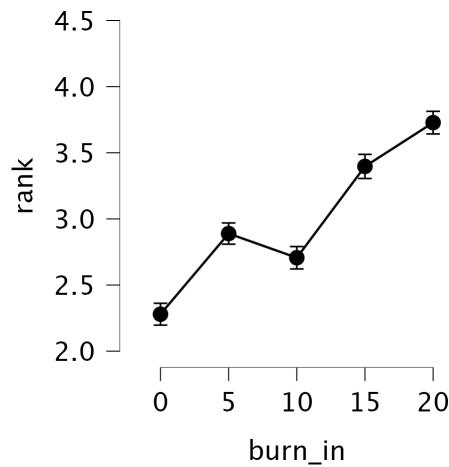
dataset: iris



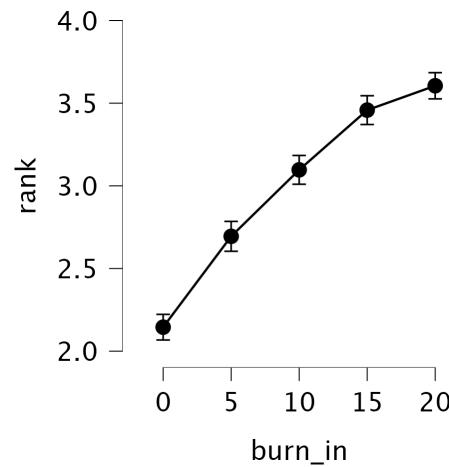
dataset: mushroom



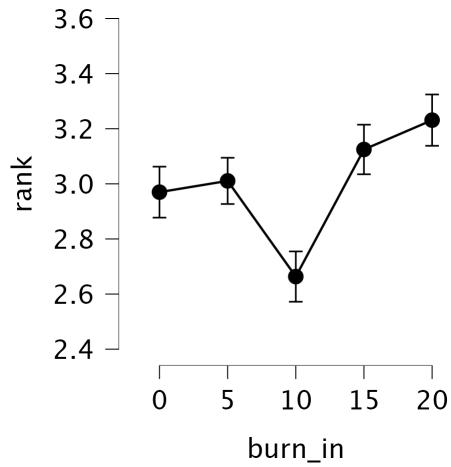
dataset: parkinsons



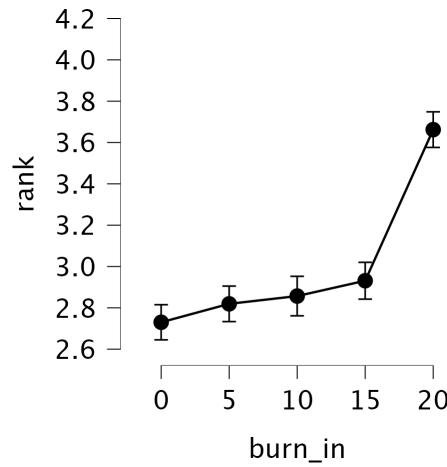
dataset: student_performance



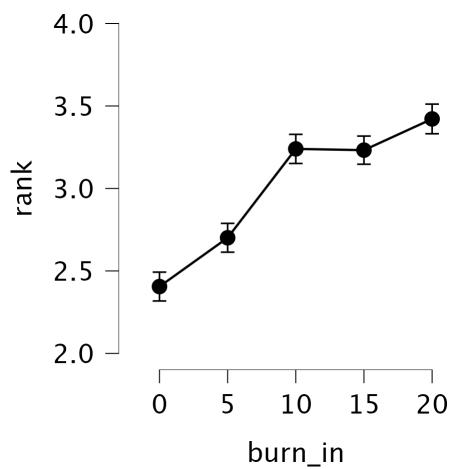
dataset: wine_quality



dataset: bank

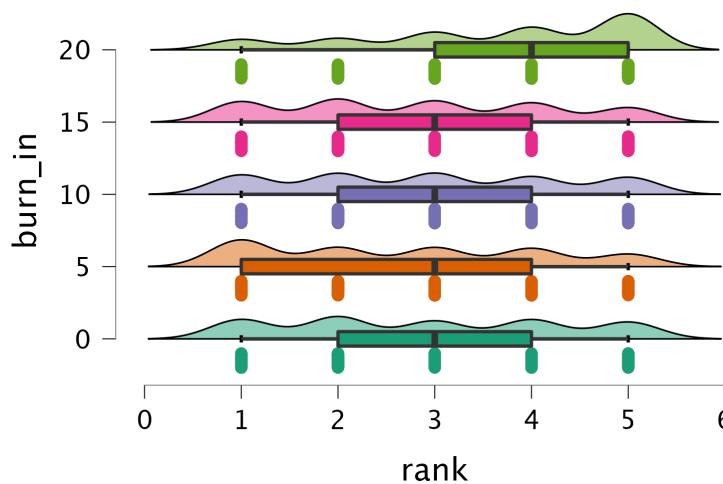


dataset: diabetic

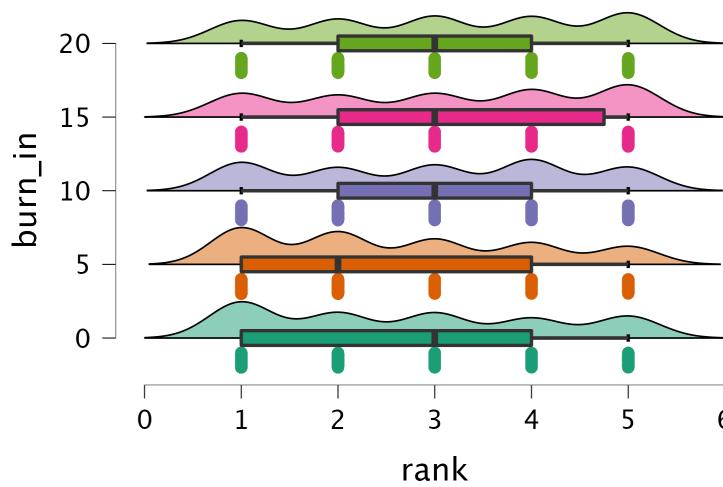


Raincloud plots

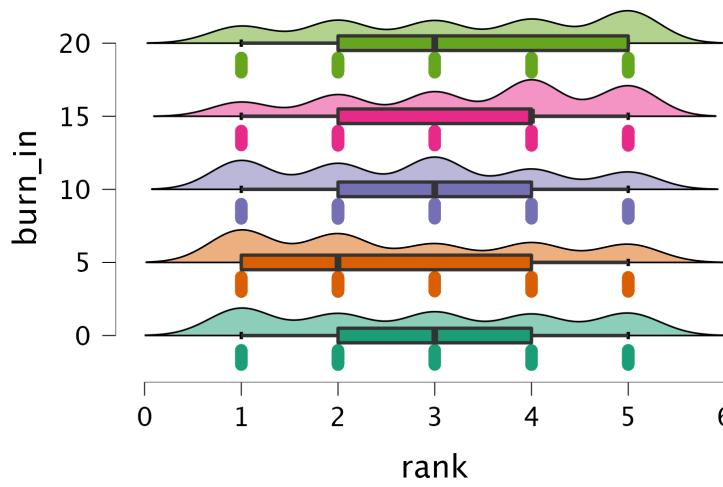
dataset: abalone



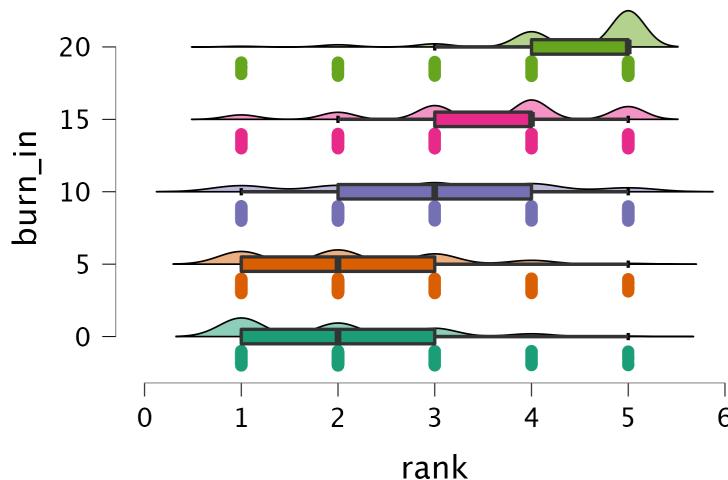
dataset: adult



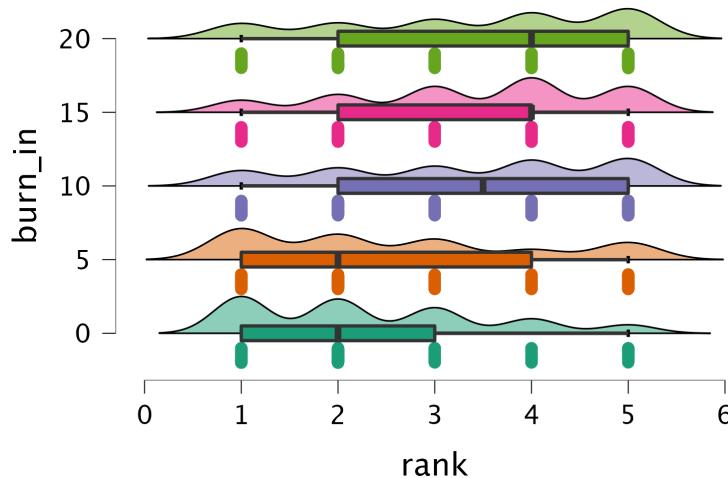
dataset: air_quality



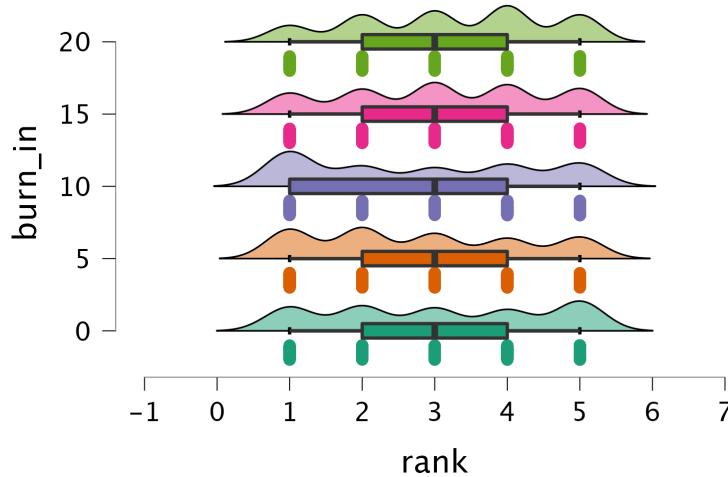
dataset: bike



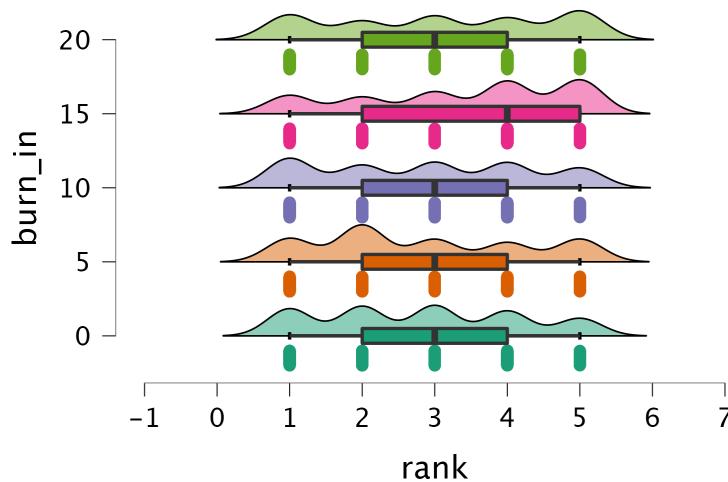
dataset: car



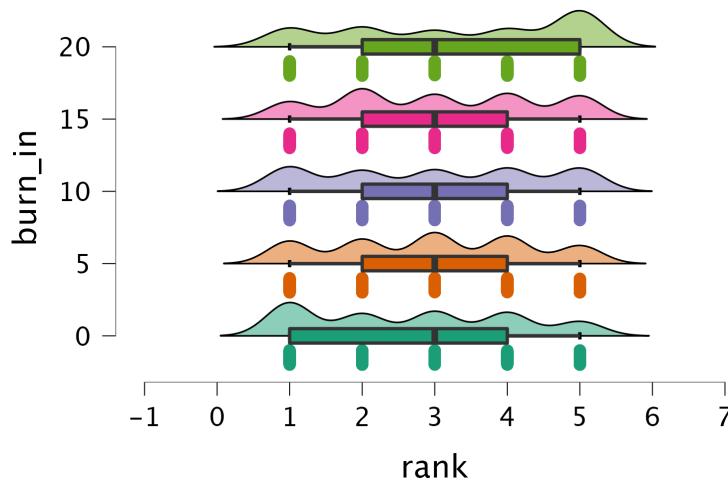
dataset: fish_toxicity



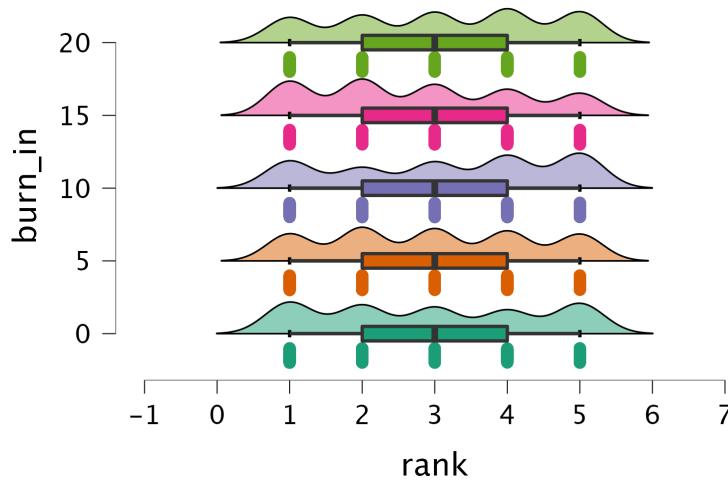
dataset: forest_fires



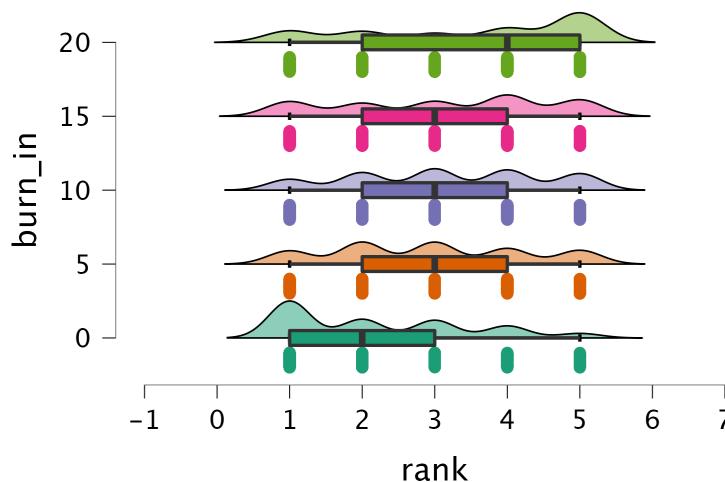
dataset: housing



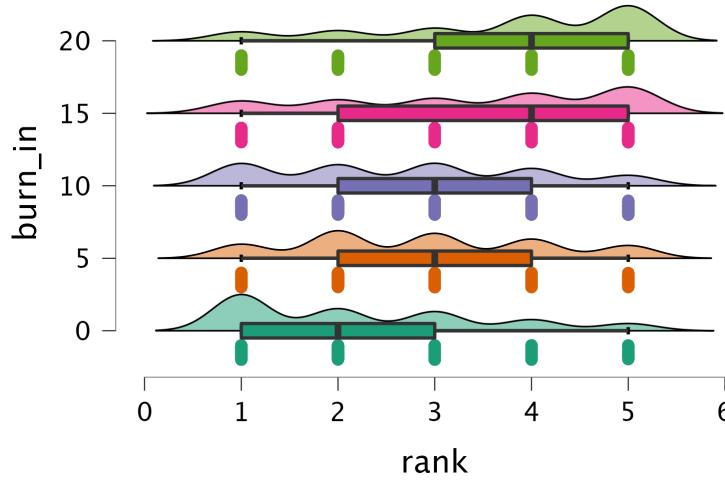
dataset: iris



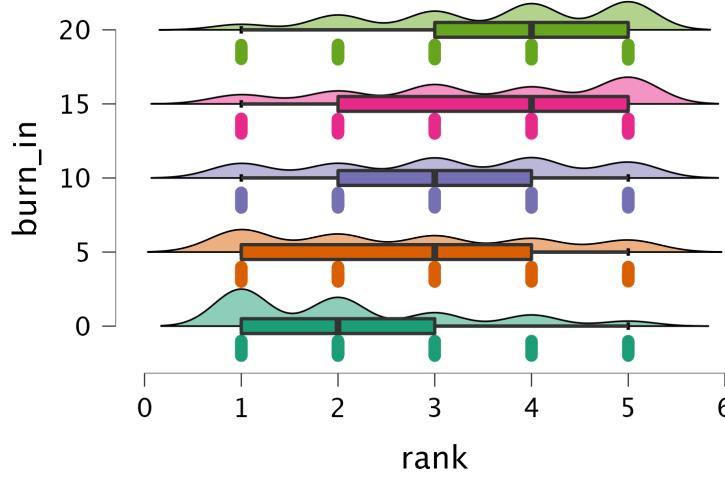
dataset: mushroom



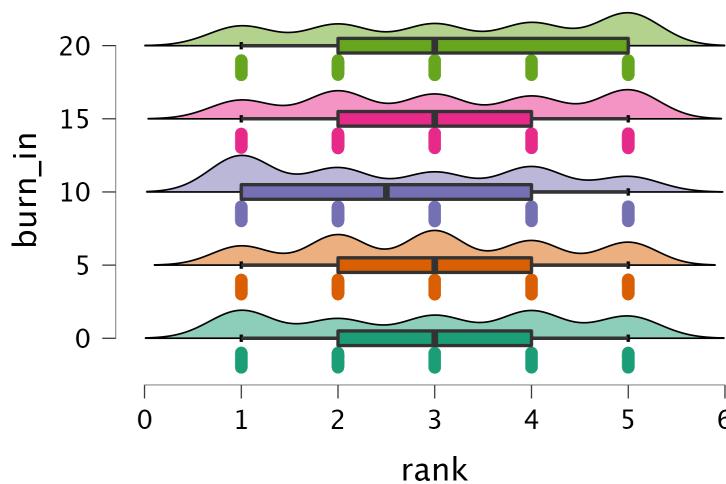
dataset: parkinsons



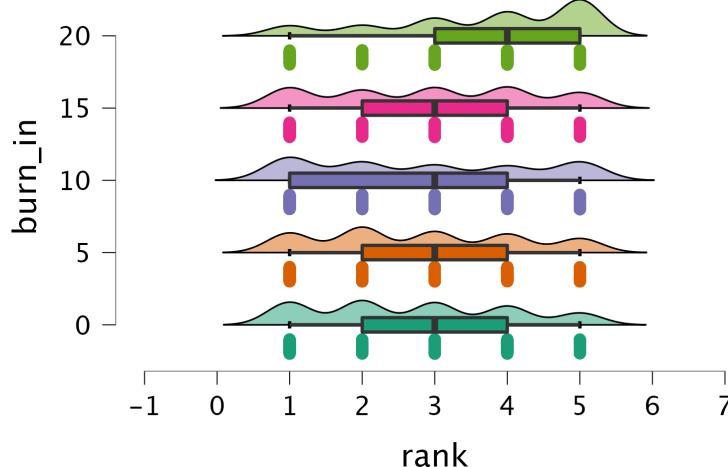
dataset: student_performance



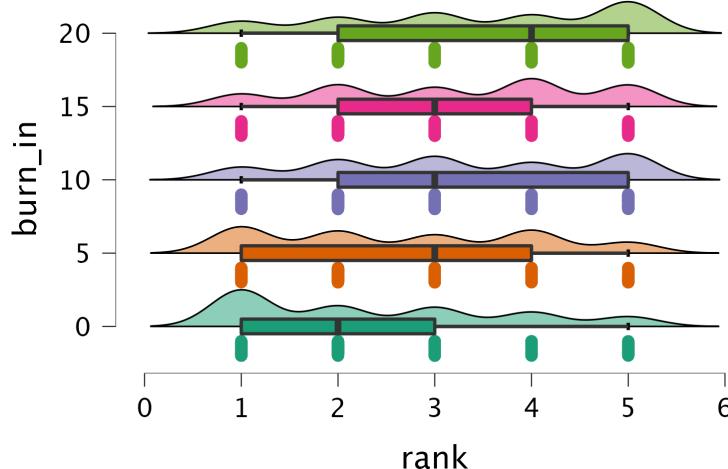
dataset: wine_quality



dataset: bank



dataset: diabetic

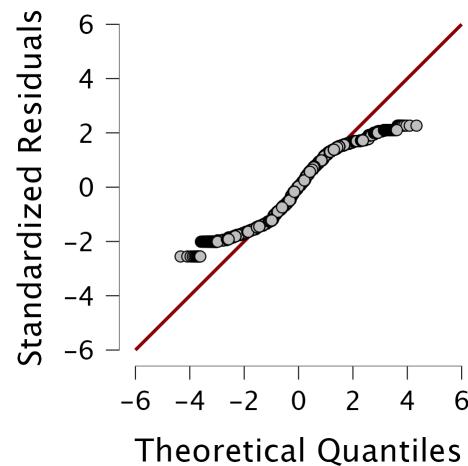


Assumption Checks

Test for Equality of Variances (Levene's)

F	df1	df2	p
25.582	74.000	69675.000	< .001

Q-Q Plot



Contrast Tables

Simple Contrast – burn_in

Comparison	Estimate	SE	df	t	p
5 - 0	0.154	0.016	69675	9.449	< .001
10 - 0	0.360	0.016	69675	22.107	< .001
15 - 0	0.584	0.016	69675	35.923	< .001
20 - 0	0.845	0.016	69675	51.975	< .001

Post Hoc Tests

Standard

Post Hoc Comparisons – burn_in

95% CI for Mean Difference						
		Mean Difference	Lower	Upper	SE	t
0	5	-0.154	-0.198	-0.109	0.016	-9.449
10	5	-0.360	-0.404	-0.315	0.016	-22.107
15	5	-0.584	-0.629	-0.540	0.016	-35.923
20	5	-0.845	-0.890	-0.801	0.016	-51.975
10	15	-0.206	-0.250	-0.162	0.016	-12.658
15	15	-0.431	-0.475	-0.386	0.016	-26.474
20	15	-0.692	-0.736	-0.647	0.016	-42.525
10	20	-0.225	-0.269	-0.180	0.016	-13.817
15	20	-0.486	-0.530	-0.441	0.016	-29.868
15	20	-0.261	-0.305	-0.217	0.016	-16.051

Note. Results are averaged over the levels of: dataset

Note. P-value and confidence intervals adjusted for comparing a family of 5 estimates (confidence intervals corrected using the tukey method).

*** p < .001

Kruskal-Wallis Test

Kruskal-Wallis Test

Factor	Statistic	df	p
burn_in	3157.390	4	< .001

