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	2.353	1.349	0.839	< .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	2.991	1.423	0.886	< .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	2.288	0.453	0.659	< .001	1.920	6.929
test_loss	adult	4650	0	242.509	705.428	0.341	< .001	0.314	20271.402
test_loss	air_quality	4650	0	0.277	0.050	0.391	< .001	0.240	0.750
test_loss	bike	4650	0	0.086	0.069	0.417	< .001	0.047	0.664
test_loss	car	4650	0	0.253	0.298	0.357	< .001	0.077	2.833
test_loss	fish_toxicity	4650	0	0.112	0.042	0.348	< .001	0.085	0.527
test_loss	forest_fires	4650	0	0.087	0.102	0.589	< .001	0.009	0.883
test_loss	housing	4650	0	0.108	0.050	0.491	< .001	0.063	0.572
test_loss	iris	4650	0	0.282	0.510	0.387	< .001	0.001	13.402
test_loss	mushroom	4650	0	2.153	127.055	0.004	< .001	0.000	8634.689
test_loss	parkinsons	4650	0	0.071	0.056	0.206	< .001	0.054	0.662
test_loss	student_performance	4650	0	0.217	0.073	0.689	< .001	0.138	0.603
test_loss	wine_quality	4650	0	1.143	0.248	0.310	< .001	0.995	2.953
test_loss	bank	4650	0	0.267	0.288	0.113	< .001	0.206	14.442
test_loss	diabetic	4650	0	1.278	1.506	0.221	< .001	0.881	46.206

ANOVA

ANOVA - rank

Cases	Sum of Squares	df	Mean Square	F	p
dataset	1814.177	14	129.584	66.508	< .001
reanalysis	97.535	4	24.384	12.515	< .001
dataset ✻ reanalysis	2924.722	56	52.227	26.805	< .001
Residuals	135755.285	69675	1.948		

Note. Type III Sum of Squares

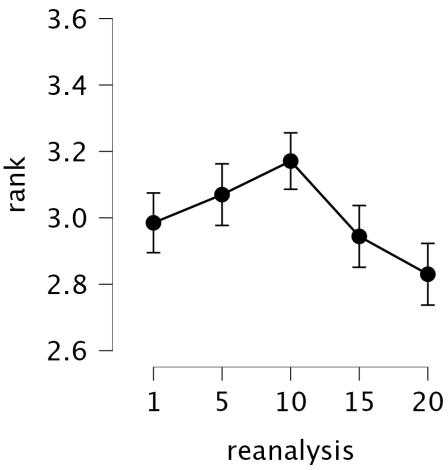
Descriptives

Descriptives - rank

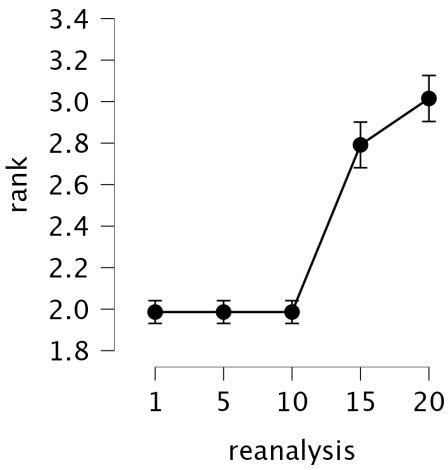
dataset	reanalysis	Mean	SD	N
abalone	1	2.985	1.397	930
	10	3.171	1.320	930
	15	2.944	1.445	930
	20	2.830	1.443	930
	5	3.070	1.442	930
adult	1	1.986	0.852	930
	10	1.986	0.852	930
	15	2.791	1.709	930
	20	3.015	1.724	930
	5	1.986	0.852	930
air_quality	1	2.961	1.344	930
	10	3.256	1.402	930
	15	3.054	1.420	930
	20	2.890	1.430	930
	5	2.839	1.438	930
bank	1	3.116	1.432	930
	10	2.735	1.398	930
	15	2.928	1.403	930
	20	3.175	1.415	930
	5	3.045	1.383	930
bike	1	3.141	1.258	930
	10	3.195	1.365	930
	15	3.098	1.498	930
	20	2.809	1.491	930
	5	2.758	1.391	930
car	1	3.124	1.423	930
	10	2.876	1.299	930
	15	3.154	1.385	930
	20	3.123	1.547	930
	5	2.724	1.358	930
diabetic	1	3.245	1.305	930
	10	2.778	1.407	930
	15	3.029	1.425	930
	20	2.875	1.494	930
	5	3.072	1.391	930
fish_toxicity	1	3.145	1.362	930
	10	3.217	1.364	930
	15	3.100	1.409	930
	20	2.534	1.388	930
	5	3.003	1.445	930
forest_fires	1	2.840	1.369	930
	10	3.004	1.318	930
	15	2.886	1.354	930
	20	3.148	1.526	930
	5	3.122	1.471	930
housing	1	3.313	1.311	930
	10	2.687	1.483	930
	15	2.697	1.496	930
	20	3.314	1.236	930
	5	2.989	1.393	930
iris	1	3.120	1.301	930
	10	3.127	1.432	930
	15	2.637	1.449	930
	20	3.019	1.435	930
	5	3.097	1.391	930
mushroom	1	2.931	1.379	930
	10	2.847	1.384	930
	15	3.122	1.453	930
	20	2.903	1.471	930
	5	3.154	1.403	930
parkinsons	1	2.969	1.385	930
	10	2.613	1.481	930
	15	3.103	1.380	930
	20	3.306	1.378	930
	5	3.009	1.356	930
student_performance	1	2.941	1.412	930
	10	2.771	1.410	930
	15	3.126	1.366	930
	20	2.977	1.467	930
	5	3.185	1.380	930
wine_quality	1	2.803	1.444	930
	10	3.202	1.410	930
	15	2.925	1.349	930
	20	3.149	1.517	930
	5	2.920	1.304	930

Descriptives plots

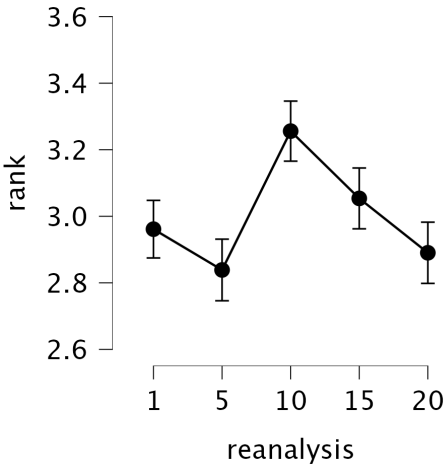
dataset: abalone



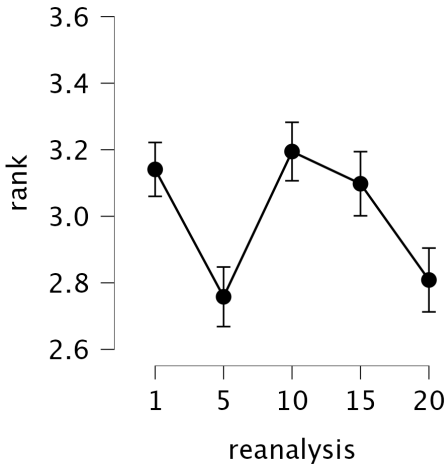
dataset: adult



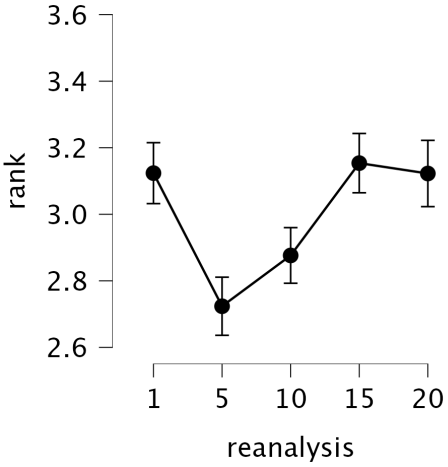
dataset: air_quality



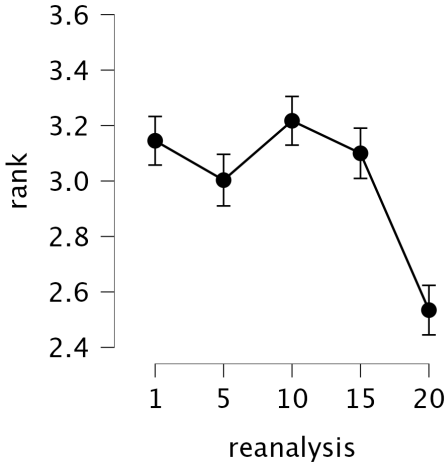
dataset: bike



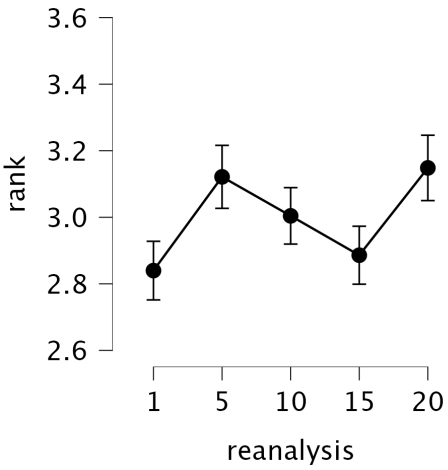
dataset: car



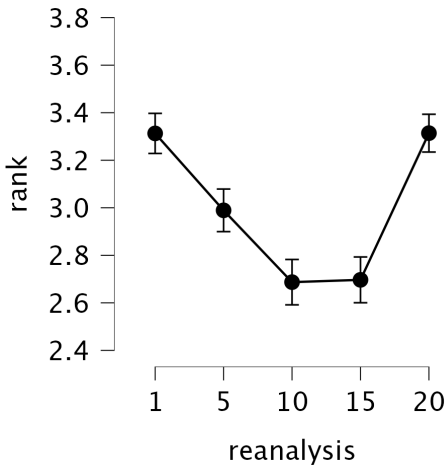
dataset: fish_toxicity



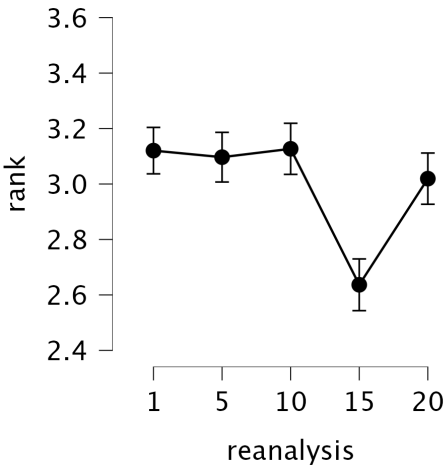
dataset: forest_fires



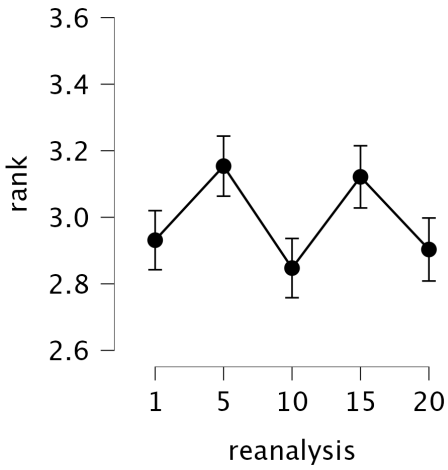
dataset: housing



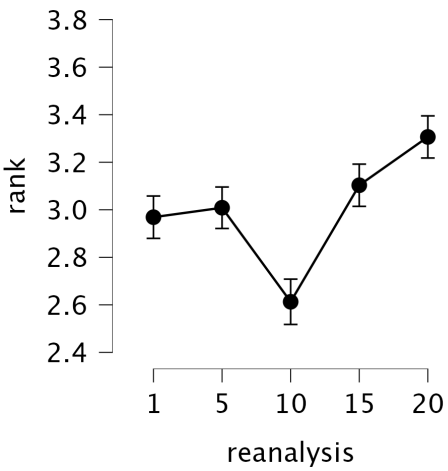
dataset: iris



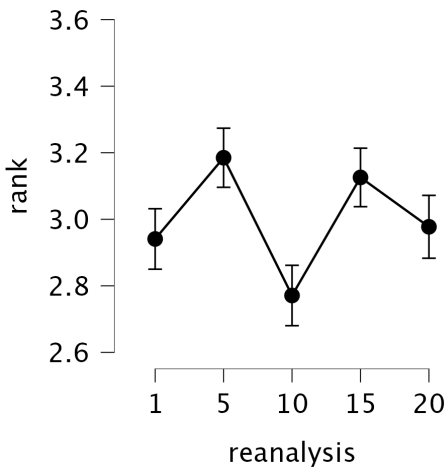
dataset: mushroom



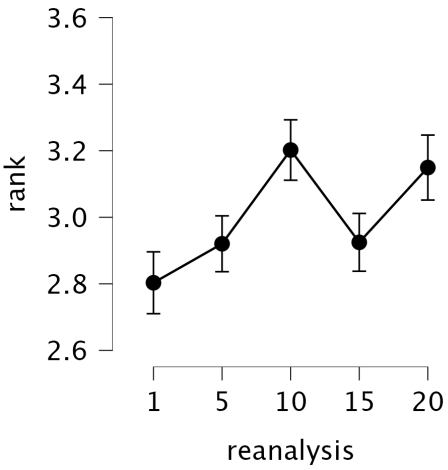
dataset: parkinsons



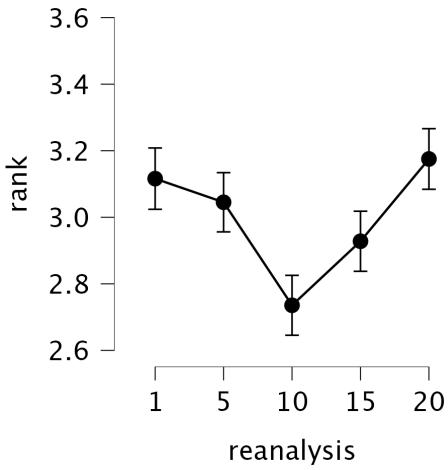
dataset: student_performance



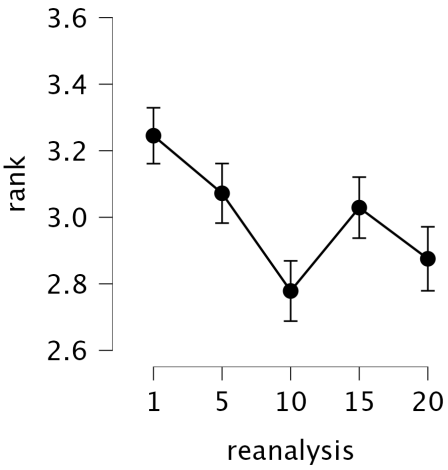
dataset: wine_quality



dataset: bank

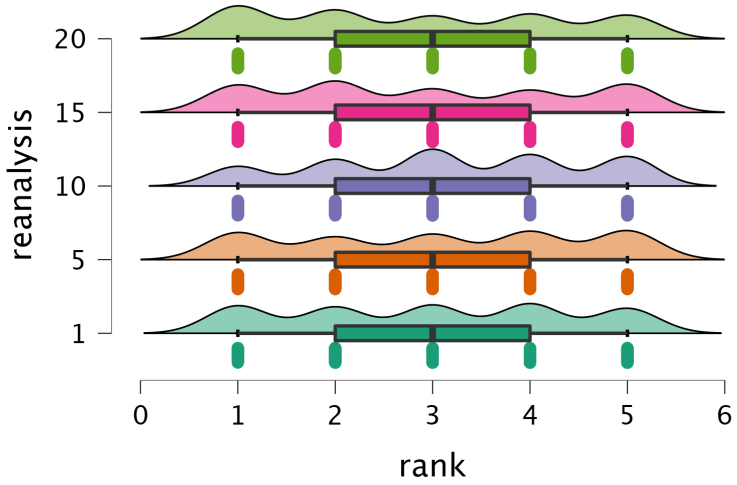


dataset: diabetic

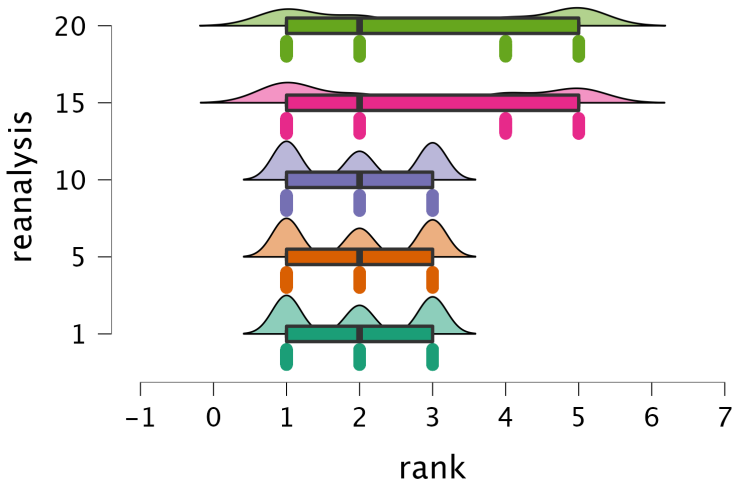


Raincloud plots

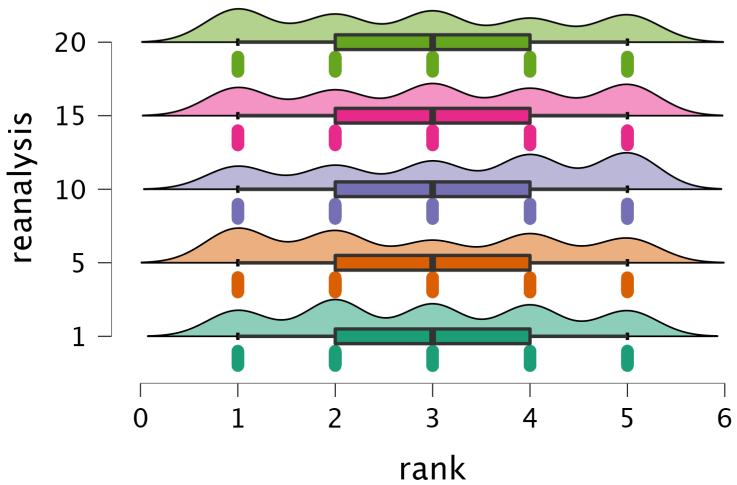
dataset: abalone



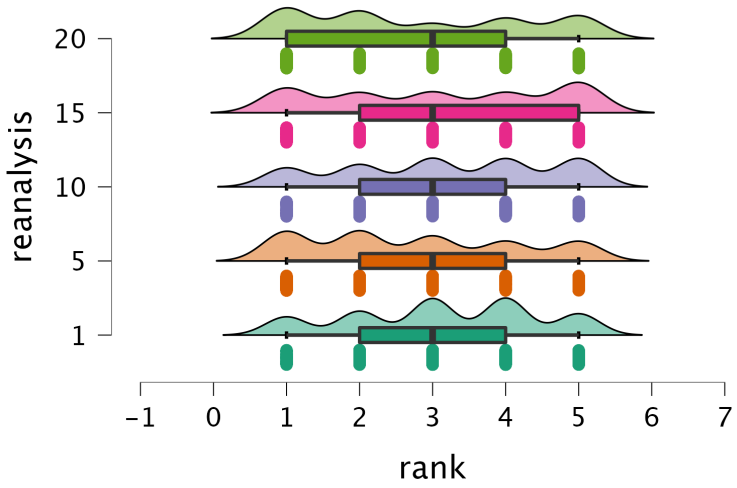
dataset: adult



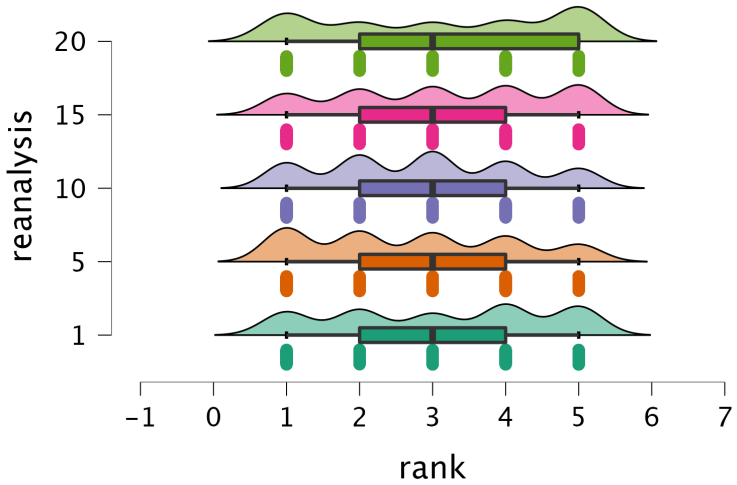
dataset: air_quality



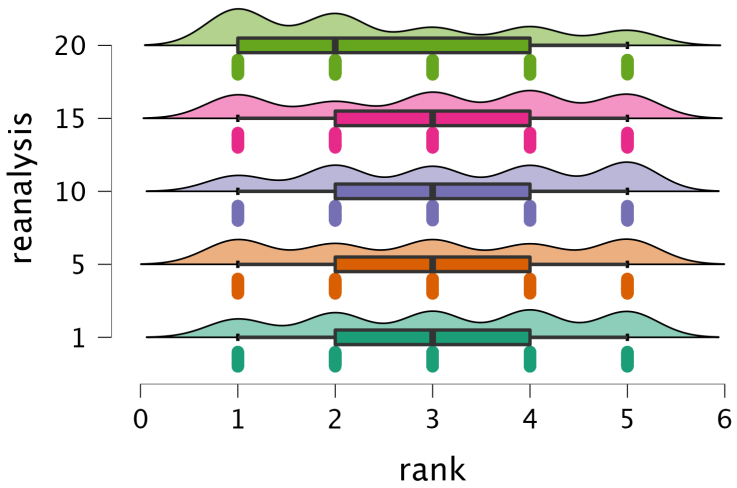
dataset: bike



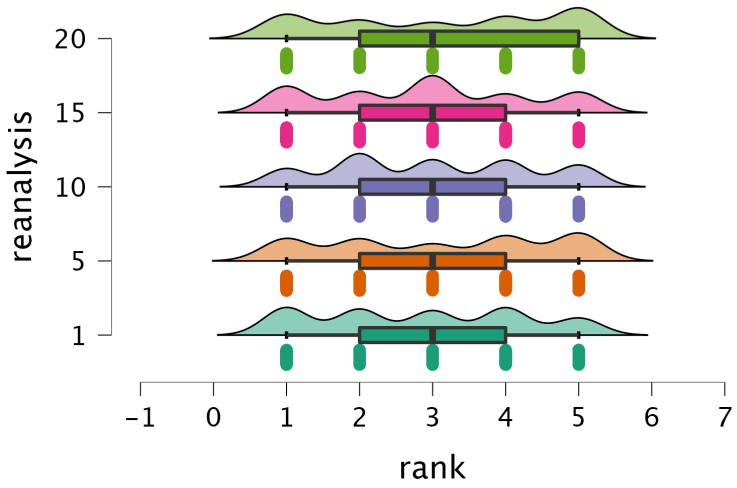
dataset: car



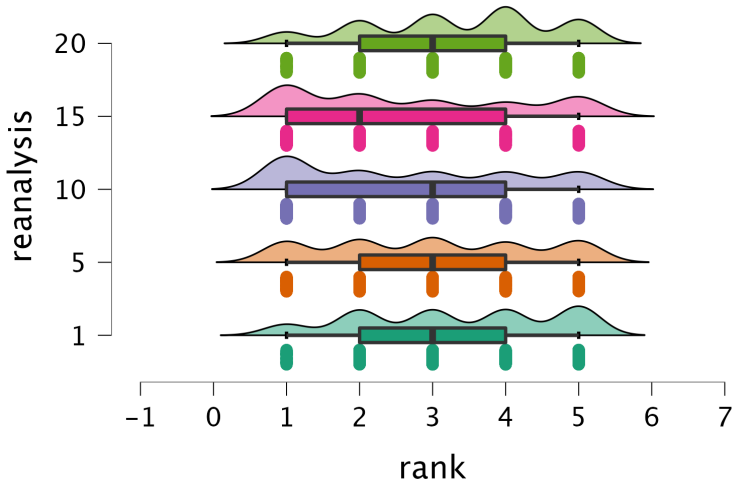
dataset: fish_toxicity



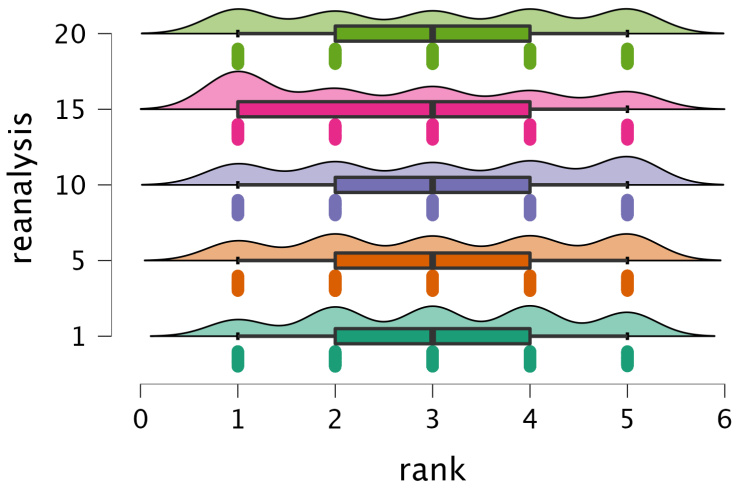
dataset: forest_fires



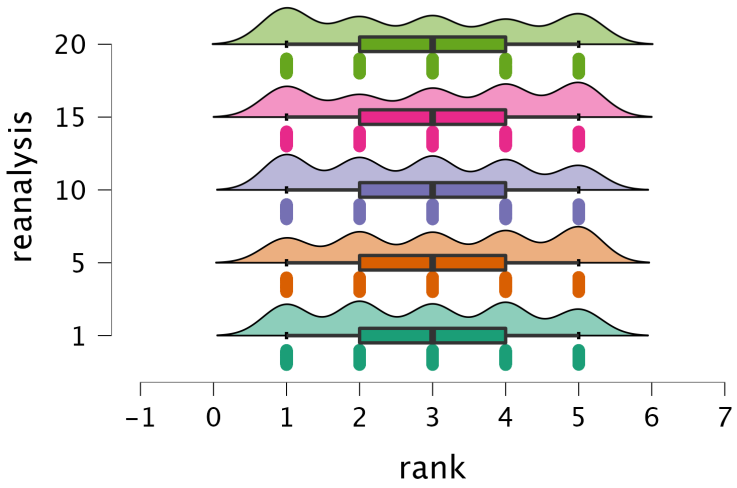
dataset: housing



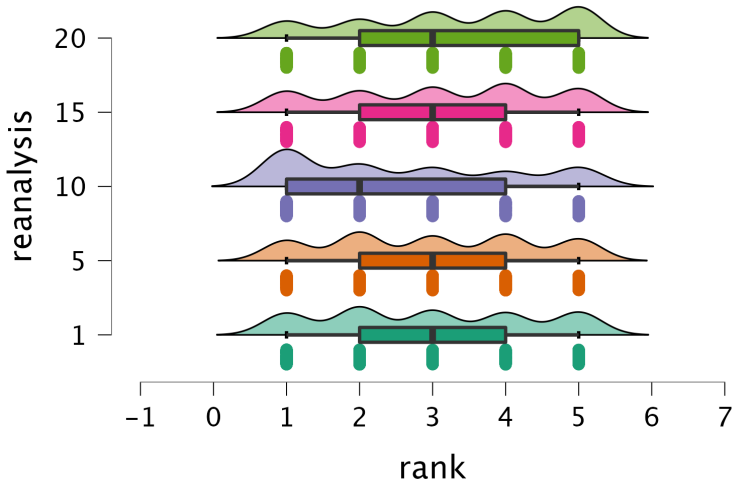
dataset: iris



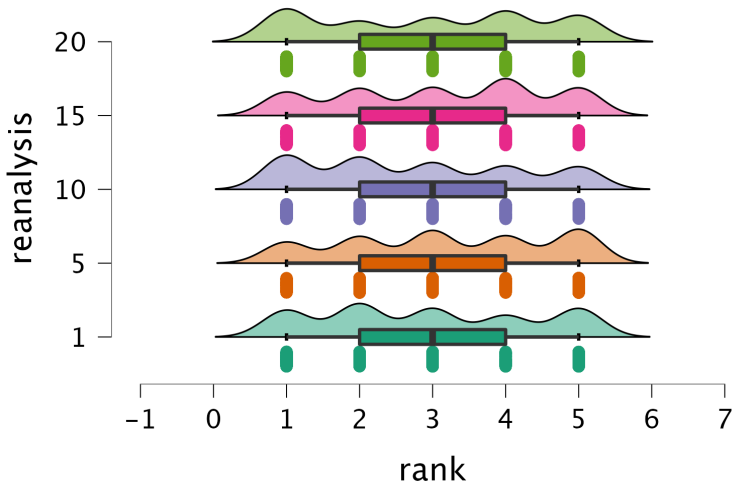
dataset: mushroom



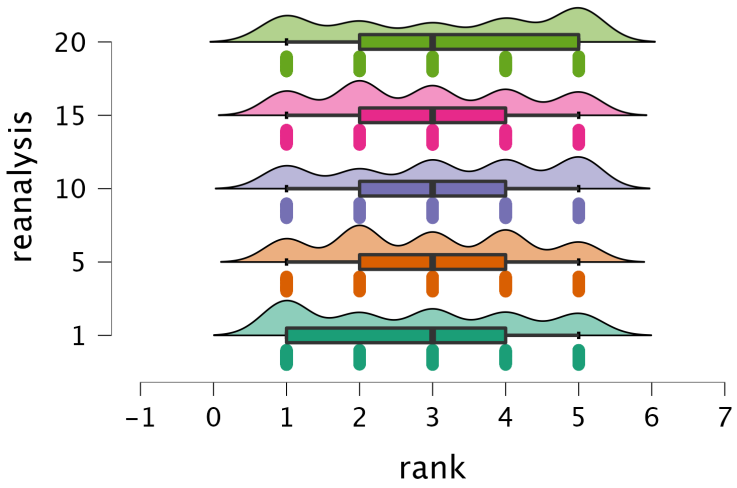
dataset: parkinsons



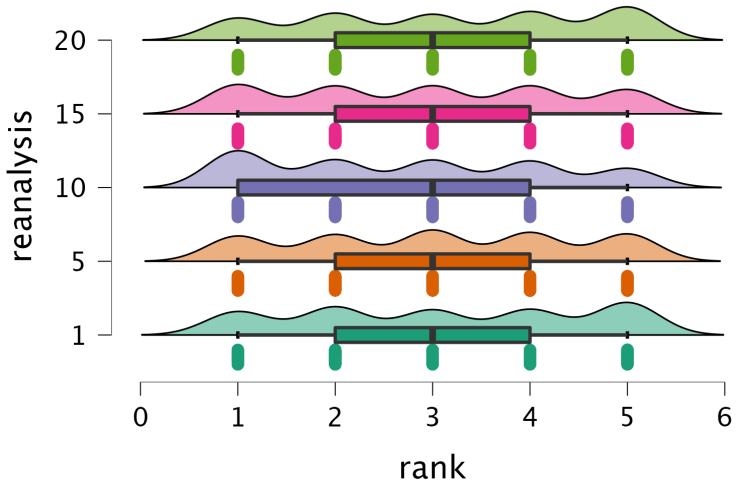
dataset: student_performance



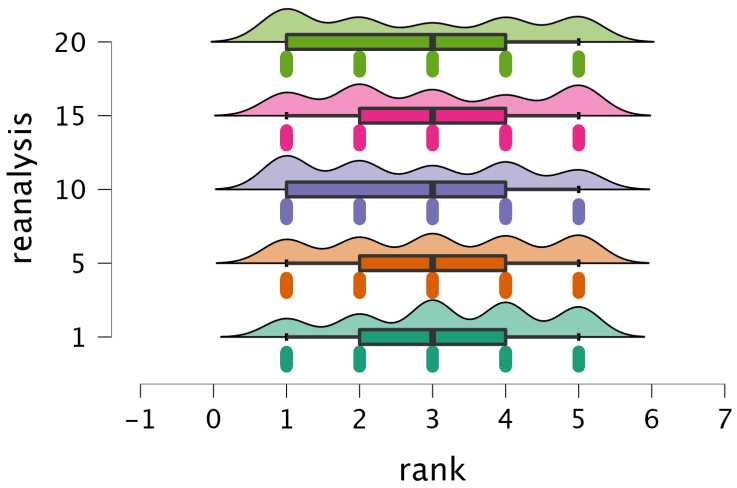
dataset: wine_quality



dataset: bank



dataset: diabetic



Assumption Checks

Test for Equality of Variances (Levene's)

F	df1	df2	p
37.312	74.000	69675.000	< .001

Contrast Tables

Simple Contrast - reanalysis

Comparison	Estimate	SE	df	t	p
5 - 1	−0.043	0.017	69675	−2.586	0.010
10 - 1	−0.077	0.017	69675	−4.602	< .001
15 - 1	−0.002	0.017	69675	−0.112	0.911
20 - 1	0.030	0.017	69675	1.793	0.073

Post Hoc Tests

Standard

Post Hoc Comparisons - reanalysis

			95% CI for Mean Difference				
		Mean Difference	Lower	Upper	SE	t	p_{tukey}
1	5	0.043	−0.002	0.089	0.017	2.586	0.073
	10	0.077	0.031	0.123	0.017	4.602	< .001***
	15	0.002	−0.044	0.047	0.017	0.112	1.000
	20	−0.030	−0.076	0.016	0.017	−1.793	0.378
5	10	0.034	−0.012	0.079	0.017	2.016	0.258
	15	−0.041	−0.087	0.004	0.017	−2.475	0.096
	20	−0.073	−0.119	−0.028	0.017	−4.379	< .001***
10	15	−0.075	−0.121	−0.029	0.017	−4.491	< .001***
	20	−0.107	−0.152	−0.061	0.017	−6.395	< .001***
15	20	−0.032	−0.077	0.014	0.017	−1.904	0.315

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
reanalysis	46.496	4	< .001