

統計グラフ作成ソフト「SGRAPPA」を 活用した統計分析

～大阪市春季総体 男子100m走を分析する～

大阪市立大正東中学校

山崎 真史

統計教育をされていて

こんなことを思ったことはありませんか？

- ▶ グラフや表を作ることで精一杯
本当は作った表やグラフからどんな傾向が読み取れるか考えさせたいのに...
- ▶ 生徒がグラフや表を作ってきたけれど、もう少し工夫すればもっと良い表やグラフを作成できるのに...
- ★ 生徒たちに与えられたデータを一人一台パソコンを使って処理させたいな
- ★ でもEXCELでは難易度が高い...

それを「S G R A P A」を使って生徒自身が分析できるようにしませんか？

昨年度は「CODAP」を使った方法を紹介しました

「SGRAPA」とは

URL: <https://sgrapa.com/>

- ▶ 株式会社正進社が開発した無料のグラフ作成ソフト
- ▶ ヒストグラムや度数分布表、ドットプロットや箱ひげ図がとても簡単に作成できる
- ▶ インターネット版、ダウンロード版がある

▶ 来年度から使用される

大日本図書株式会社

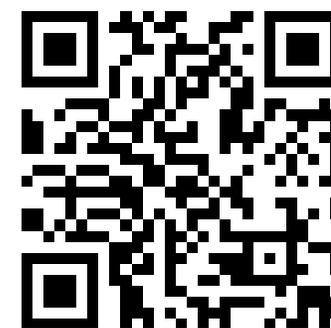
日本文教出版株式会社

学校図書株式会社

教育出版株式会社

の計4社の教科書にて使用される予定

SGRAPA - 無料で統計



グラフが作成できる

大阪市春季総合体育大会100m走を分析し、 どのくらいのタイムを出せば府大会に進出できるのか 考えよう

▶ 予想する

- ・ 2023年度の大会の決勝進出ラインは12.02秒 ・ 年々早くなっている？

▶ 集める

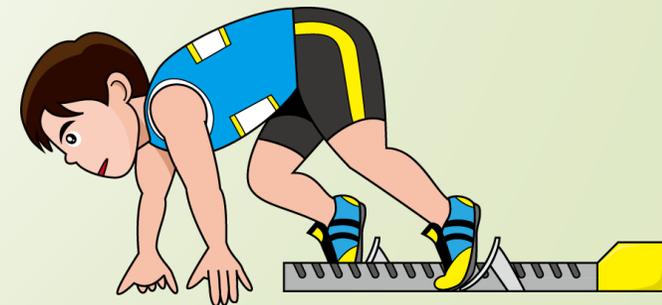
- ・ 大阪市中体連陸上競技部のホームページにあるリザルト集から、2016年から2024年の大阪市春季総体の男子100mの結果を集め、表計算ソフトに情報を入力する。

▶ 分析する

- ・ 入力したデータをタイムの速い順に並べ替え、度数分布表や箱ひげ図をつくり、傾向を調べる。(府大会進出は18位まで)

▶ まとめる

- ・ 傾向を分析できたら、府大会進出のための目標タイムを設定し、日々トレーニングをする。



データを処理し、分析する

	A	B	C	D	E	F	G	H	I	J
1		2024	2023	2022	2021	2020	2019	2018	2017	2016
2		記録								
3	1	11.11	11.42	11.19	11.28	11.12	11.48	11.64	11.44	11.37
4	2	11.51	11.47	11.59	11.54	11.38	11.51	11.67	11.47	11.47
5	3	11.53	11.60	11.66	11.57	11.55	11.55	11.71	11.50	11.49
6	4	11.57	11.64	11.67	11.58	11.77	11.72	11.86	11.54	11.62
7	5	11.62	11.66	11.67	11.63	11.82	11.81	11.90	11.60	11.71
8	6	11.67	11.70	11.72	11.69	11.86	11.82	11.91	11.67	11.75
9	7	11.76	11.73	11.74	11.71	11.93	11.82	11.92	11.73	11.77
10	8	11.81	11.74	11.77	11.77	11.96	11.82	11.93	11.75	11.80
11	9	11.83	11.75	11.81	11.81	11.98	11.85	11.95	11.75	11.92
12	10	11.85	11.76	11.83	11.82	12.04	11.87	11.95	11.80	12.00
13	11	11.85	11.78	11.86	11.83	12.05	11.93	11.98	11.82	12.01
14	12	11.86	11.85	11.88	11.90	12.09	11.94	11.98	11.85	12.02
15	13	11.94	11.88	11.88	11.92	12.12	11.95	12.04	11.89	12.08
16	14	11.96	11.93	11.88	11.93	12.15	11.97	12.05	11.90	12.12
17	15	12.00	11.94	11.89	11.97	12.16	12.01	12.08	11.91	12.12
18	16	12.00	11.95	11.92	11.98	12.16	12.04	12.10	11.94	12.16
19	17	12.02	11.97	11.95	12.00	12.18	12.09	12.17	11.95	12.17
20	18	12.02	12.01	11.96	12.01	12.21	12.12	12.21	11.99	12.18
21	19	12.02	12.04	11.96	12.01	12.27	12.15	12.21	12.00	12.19
22	20	12.04	12.06	11.97	12.05	12.30	12.15	12.21	12.04	12.19

タイム順に並べ替えた結果

・府大会に進出できる18番目までに入ろうと思ったら、12.00秒前後までに走らないといけない。

・決勝(8番)進出するためには、11.80秒ぐらいのタイムを出さないといけない。

疑問

・2020年、2018年、2016年の決勝進出ラインはほかの年に比べてタイムが遅い。何がいった？

全体の記録を見返し、分析してみよう

データを処理する(度数分布表)

The screenshot shows the SGRAPA software interface. The top left has the SGRAPA logo and a project name field. Below it is a 'データ入力' (Data Input) section with a grid of data. The grid has columns labeled A through Q and rows numbered 1 through 31. A red box highlights the data range from row 2 to row 31, columns A to I. To the right of the grid is a 'グラフ表示' (Graph Display) section with a '100%' zoom level. At the bottom, there is a navigation bar with options: 'ドットプロット', '度数分布表', 'ヒストグラム', and '箱ひげ図'. The '度数分布表' option is circled in red. Two numbered instructions are overlaid on the grid: ① データの範囲を選択し (Select the data range) and ② 「度数分布表」をクリックするだけで (Clicking 'Frequency Distribution Table' is enough).

SGRAPA プロジェクト名 中2・3モード

データ入力

数式

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1																	
2	2024年	2023年	2022年	2021年	2020年	2019年	2018年	2017年	2016年								
3	11.11	11.42	11.19	11.28	11.12	11.48	11.64	11.44	11.37								
4	11.51	11.47	11.59	11.54	11.38	11.51	11.67	11.47	11.47								
5	11.53	11.6	11.66	11.57	11.55	11.55	11.71	11.5	11.49								
6	11.57	11.64	11.67	11.58	11.77	11.72	11.86	11.54	11.62								
7	11.62	11.66	11.67	11.63	11.82	11.81	11.9	11.6	11.71								
8	11.67	11.7	11.72	11.69	11.86	11.82	11.91	11.67	11.75								
9	11.76	11.73	11.74	11.71	11.93	11.82	11.92	11.73	11.77								
10	11.81	11.74	11.77	11.77	11.96	11.82	11.93	11.75	11.8								
11	11.83	11.75	11.81	11.81	11.98	11.85	11.95	11.75	11.92								
12	11.85	11.76	11.83	11.82	12.04	11.87	11.95	11.8	12								
13	11.85	11.78	11.86	11.83	12.05	11.93	11.98	11.82	12.01								
14	11.86	11.85	11.88	11.9	12.09	11.94	11.98	11.85	12.02								
15	11.94	11.88	11.88	11.92	12.12	11.95	12.04	11.89	12.08								
16	11.96	11.93	11.88	11.93	12.15	11.97	12.05	11.9	12.12								
17	12	11.94	11.89	11.97	12.16	12.01	12.08	11.91	12.12								
18	12	11.95	11.92	11.98	12.16	12.04	12.1	11.94	12.16								
19	12.02	11.97	11.95	12	12.18	12.09	12.17	11.95	12.17								
20	12.02	12.01	11.96	12.01	12.21	12.12	12.21	11.99	12.18								
21	12.02	12.04	11.96	12.01	12.27	12.15	12.21	12	12.19								
22	12.04	12.06	11.97	12.05	12.3	12.15	12.21	12.04	12.19								
23	12.05	12.1	11.98	12.08	12.3	12.15	12.24	12.04	12.2								
24	12.05	12.13	12.02	12.13	12.37	12.17	12.24	12.08	12.24								
25	12.08	12.15	12.02	12.19	12.39	12.18	12.31	12.13	12.27								
26	12.09	12.18	12.03	12.25	12.4	12.2	12.33	12.14	12.35								
27	12.11	12.18	12.04	12.25	12.41	12.24	12.35	12.15	12.36								
28	12.11	12.19	12.04	12.26	12.43	12.24	12.39	12.15	12.38								
29	12.15	12.19	12.06	12.29	12.45	12.25	12.4	12.16	12.41								
30	12.15	12.2	12.09	12.3	12.45	12.25	12.44	12.18	12.42								
31	12.16	12.2	12.1	12.35	12.47	12.25	12.46	12.2	12.43								

① データの範囲を選択し

② 「度数分布表」をクリックするだけで

ドットプロット 度数分布表 ヒストグラム 箱ひげ図

データを処理する(度数分布表)

SGRAPA プロジェクト名 中2・3モード

データ入力

行数	列	内容
136	DNS	DNS
137	DNS	DNS
138	DNS	DNS
139	DNS	DNS
140	DNS	DNS
141	DNS	DNS
142	DNS	DNS
143	DNS	DNS
144	DNS	DNS
145	DNS	DNS
146	DNS	DNS
147	DNS	DNS
148	DNS	DNS
149	DNS	DNS
150	DNS	DNS
151	DNS	DNS
152	DNS	DNS
153	DNS	DNS
154	DNS	DNS
155	DNS	DNS
156	DNS	DNS
157	DNS	DNS
158	DNS	DNS
159	DNS	DNS
160		
161		
162		
163		
164		
165		
166		
167		
168		
169		
170		

グラフ表示 100%

階級	2024年	2023年	2022年	2021年	2020年	2019年	2018年	2017年	2016年
	度数								
以上									
11.11 ~ 12.74	64	58	79	61	50	57	55	65	51
12.74 ~ 14.37	53	60	58	66	52	71	69	68	78
14.37 ~ 16.00	9	9	10	6	7	17	13	11	14
16.00 ~ 17.63	5	1	1	3	1	1	3	1	6
17.63 ~ 19.26	0	0	0	0	0	0	0	0	0
19.26 ~ 20.89	1	0	0	0	0	0	0	0	0
20.89 ~ 22.52	0	1	0	0	0	0	0	0	0
22.52 ~ 24.15	0	0	0	0	1	0	0	0	0
合計	132	129	148	136	111	146	140	145	149

③ 簡単に度数分布表を作ることができる

④ 次にここをクリックし、階級の幅を調整し、見やすい度数分布表へ

分析する(度数分布表)

SGRAPA プロジェクト名 中2・3モード

データ入力

数式	DNS																		
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
136	DNS	DNS	14.26	16.1		14.79	15	14.43	14.77										
137	DNS	DNS	14.28	16.8		14.86	15.04	14.44	14.78										
138	DNS	DNS	14.28	16.87		14.9	15.37	14.78	14.81										
139	DNS	DNS	14.31	DNS		14.93	15.73	14.84	14.83										
140	DNS	DNS	14.65	DNS		15.03	16.19	15.1	14.83										
141	DNS	DNS	14.71	DNS		15.05	16.21	15.11	15.4										
142	DNS	DNS	14.83	DNS		15.11	16.28	15.34	15.61										
143	DNS	DNS	14.88	DNS		15.18	DNS	15.38	15.66										
144	DNS	DNS	14.92	DNS		15.34	DNS	15.5	15.68										
145	DNS	DNS	15.05	DNS		15.45	DNS	15.56	15.81										
146	DNS	DNS	15.12	DNS		15.77	DNS	15.63	16.13										
147	DNS	DNS	15.48	DNS		15.87	DNS	16.86	16.24										
148	DNS	DNS	15.52	DNS		16.05	DNS	DNS	16.44										
149	DNS	DNS	15.96	DNS		DNS	DNS	DNS	16.64										
150	DNS	DNS	16.51	DNS		DNS	DNS	DNS	17.32										
151	DNS	DNS	DNF	DNS		DNS	DNS	DNS	17.45										
152	DNS	DNS	DNS	NM		DNS	DNS	DNS	DNS										
153	DNS	DNS	DNS			DNS	DNS	DNS	DNS										
154	DNS	DNS	DNS			DNS	DNS	DNS	DNS										
155	DNS	DNS	DNS			DNS	NM	DNS	DNS										
156	DNS	NM	DNS			DNS	NM	DNS	DNS										
157	DNS		DNS			DNS	検作減	DNS	DNS										
158	DNS		DNS			DNS	検作減	DNS	DNS										
159	DNS		DNS			DNS		DNS	DNS										
160			DNS			NM		NM	検作減										
161			DNS					NM	検作減										
162								NM	検作減										
163								検作減											
164								検作減											
165								検作減											

⑤ 見やすい度数分布表の完成

グラフ表示 100%

階級	2024年	2023年	2022年	2021年	2020年	2019年	2018年	2017年	2016年
	度数								
以上									
11.00 ~ 11.50	1	2	1	1	2	1	0	2	3
11.50 ~ 12.00	13	15	20	15	7	13	12	16	6
12.00 ~ 12.50	31	21	38	23	21	31	18	32	24
12.50 ~ 13.00	36	40	37	44	34	30	45	29	36
13.00 ~ 13.50	17	18	26	29	27	30	22	25	36
13.50 ~ 14.00	16	15	10	11	9	19	14	18	13
14.00 ~ 14.50	4	8	5	5	3	8	18	13	11
14.50 ~ 15.00	5	3	5	2	4	5	4	2	9
15.00 ~ 15.50	2	3	3	2	2	6	3	4	1
15.50 ~ 16.00	1	2	2	1	0	2	1	3	4
16.00 ~ 16.50	1	1	0	1	0	1	3	0	3
16.50 ~ 17.00	2	0	1	2	1	0	0	1	1
17.00 ~ 17.50	2	0	0	0	0	0	0	0	2
17.50 ~ 18.00	0	0	0	0	0	0	0	0	0
18.00 ~ 18.50	0	0	0	0	0	0	0	0	0
18.50 ~ 19.00	0	0	0	0	0	0	0	0	0
19.00 ~ 19.50	0	0	0	0	0	0	0	0	0
19.50 ~ 20.00	1	0	0	0	0	0	0	0	0
20.00 ~ 20.50	0	0	0	0	0	0	0	0	0
20.50 ~ 21.00	0	1	0	0	0	0	0	0	0
21.00 ~ 21.50	0	0	0	0	0	0	0	0	0
21.50 ~ 22.00	0	0	0	0	0	0	0	0	0

度数分布表

データを処理し、分析する

・2020年
極端に走った選手が少ないのは？

新型コロナの影響で
参加校が少なかった

階級 (秒)	2024年	2023年	2022年	2021年	2020年	2019年	2018年	2017年	2016年
	度数 (人)								
以上 未満									
11.00 ~ 11.50	1	2	1	1	2	1	0	2	3
11.50 ~ 12.00	13	15	20	15	7	13	12	16	6
12.00 ~ 12.50	31	21	38	23	21	31	18	32	24
12.50 ~ 13.00	36	40	37	44	34	30	45	29	36
13.00 ~ 13.50	17	18	26	29	27	30	22	25	36
13.50 ~ 14.00	16	15	10	11	9	19	14	18	13
14.00 ~ 14.50	4	8	5	5	3	8	18	13	11
14.50 ~ 15.00	5	3	5	2	4	5	4	2	9
15.00 ~ 15.50	2	3	3	2	2	6	3	4	1
15.50 ~ 16.00	1	2	2	1	0	2	1	3	4
16.00 ~ 16.50	1	1	0	1	0	1	3	0	3
16.50 ~ 17.00	2	0	1	2	1	0	0	1	1
17.00 ~ 17.50	2	0	0	0	0	0	0	0	2
17.50 ~ 18.00	0	0	0	0	0	0	0	0	0
18.00 ~ 18.50	0	0	0	0	0	0	0	0	0
18.50 ~ 19.00	0	0	0	0	0	0	0	0	0
19.00 ~ 19.50	0	0	0	0	0	0	0	0	0
19.50 ~ 20.00	1	0	0	0	0	0	0	0	0
20.00 ~ 20.50	0	0	0	0	0	0	0	0	0
20.50 ~ 21.00	0	1	0	0	0	0	0	0	0
21.00 ~ 21.50	0	0	0	0	0	0	0	0	0
21.50 ~ 22.00	0	0	0	0	0	0	0	0	0
22.00 ~ 22.50	0	0	0	0	0	0	0	0	0
22.50 ~ 23.00	0	0	0	0	0	0	0	0	0
23.00 ~ 23.50	0	0	0	0	0	0	0	0	0
23.50 ~ 24.00	0	0	0	0	0	0	0	0	0
24.00 ~ 24.50	0	0	0	0	1	0	0	0	0
合計	132	129	148	136	111	146	140	145	149

複数のデータを効率よく比較しよう

箱ひげ図

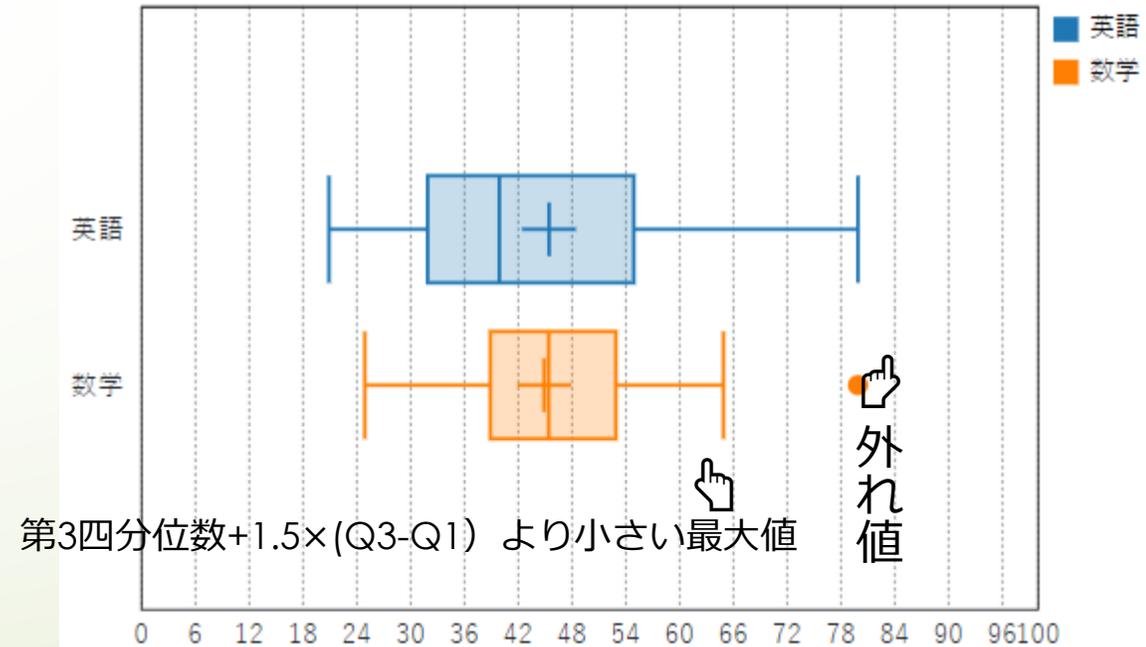
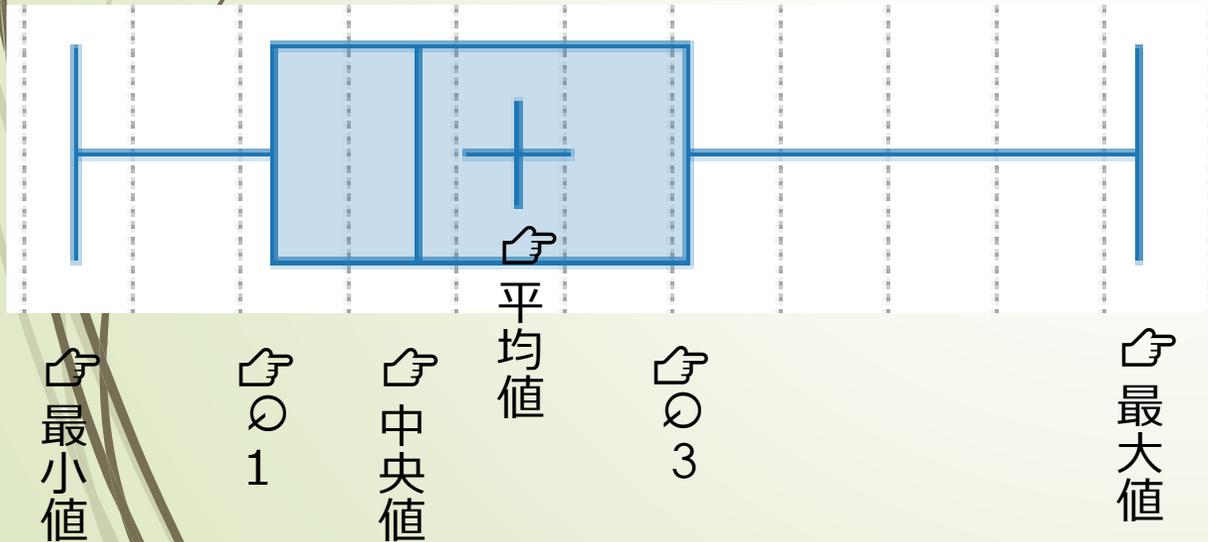
例 11人の英語と数学のテストの結果

英語	21, 29, 32, 36, 38, 40, 49, 53, 55, 68, 80
数学	25, 31, 39, 42, 45, 46, 50, 53, 54, 65, 80

第1四分位数：Q1

第2四分位数：Q2

第3四分位数：Q3



データを処理する(箱ひげ図)

The screenshot shows the SGRAPA software interface. At the top, there is a header with the SGRAPA logo, a project name field, and a mode selector set to '中2・3モード'. Below the header is a toolbar with various icons for data input and chart manipulation. The main area is divided into two panes. The left pane is a spreadsheet with columns labeled A through Q and rows numbered 1 through 31. A red box highlights a range of data from row 2 to row 31, columns A through I. The right pane is currently empty, with a 'グラフ表示' (Graph Display) button and a zoom level of 100%. At the bottom, a menu bar contains several options: 'ドットプロット', '度数分布表', 'ヒストグラム', and '箱ひげ図'. The '箱ひげ図' option is circled in red, and a red arrow points from this menu item to the highlighted data range in the spreadsheet.

SGRAPA プロジェクト名 中2・3モード

データ入力

数式

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1																	
2	2024年	2023年	2022年	2021年	2020年	2019年	2018年	2017年	2016年								
3	11.11	11.42	11.19	11.28	11.12	11.48	11.64	11.44	11.37								
4	11.51	11.47	11.59	11.54	11.38	11.51	11.67	11.47	11.47								
5	11.53	11.6	11.66	11.57	11.55	11.55	11.71	11.5	11.49								
6	11.57	11.64	11.67	11.58	11.77	11.72	11.86	11.54	11.62								
7	11.62	11.66	11.67	11.63	11.82	11.81	11.9	11.6	11.71								
8	11.67	11.7	11.72	11.69	11.86	11.82	11.91	11.67	11.75								
9	11.76	11.73	11.74	11.71	11.93	11.82	11.92	11.73	11.77								
10	11.81	11.74	11.77	11.77	11.96	11.82	11.93	11.75	11.8								
11	11.83	11.75	11.81	11.81	11.98	11.85	11.95	11.75	11.92								
12	11.85	11.76	11.83	11.82	12.04	11.87	11.95	11.8	12								
13	11.85	11.78	11.86	11.83	12.05	11.93	11.98	11.82	12.01								
14	11.86	11.85	11.88	11.9	12.09	11.94	11.98	11.85	12.02								
15	11.94	11.88	11.88	11.92	12.12	11.95	12.04	11.89	12.08								
16	11.96	11.93	11.88	11.93	12.15	11.97	12.05	11.9	12.12								
17	12	11.94	11.89	11.97	12.16	12.01	12.08	11.91	12.12								
18	12	11.95	11.92	11.98	12.16	12.04	12.1	11.94	12.16								
19	12.02	11.97	11.95	12	12.18	12.09	12.17	11.95	12.17								
20	12.02	12.01	11.96	12.01	12.21	12.12	12.21	11.99	12.18								
21	12.02	12.04	11.96	12.01	12.27	12.15	12.21	12	12.19								
22	12.04	12.06	11.97	12.05	12.3	12.15	12.21	12.04	12.19								
23	12.05	12.1	11.98	12.08	12.3	12.15	12.24	12.04	12.2								
24	12.05	12.13	12.02	12.13	12.37	12.17	12.24	12.08	12.24								
25	12.08	12.15	12.02	12.19	12.39	12.18	12.31	12.13	12.27								
26	12.09	12.18	12.03	12.25	12.4	12.2	12.33	12.14	12.35								
27	12.11	12.18	12.04	12.25	12.41	12.24	12.35	12.15	12.36								
28	12.11	12.19	12.04	12.26	12.43	12.24	12.39	12.15	12.38								
29	12.15	12.19	12.06	12.29	12.45	12.25	12.4	12.16	12.41								
30	12.15	12.2	12.09	12.3	12.45	12.25	12.44	12.18	12.42								
31	12.16	12.2	12.1	12.35	12.47	12.25	12.46	12.2	12.43								

①データの範囲を選択し

②「箱ひげ図」をクリックするだけで

ドットプロット 度数分布表 ヒストグラム 箱ひげ図

データを処理する(箱ひげ図)

SGRAPA 大阪市春季総体 中2・3モード

データ入力 2016年

数式	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
138	DNS	DNS	14.28	16.07		14.9	15.31	14.78	14.81								
139	DNS	DNS	14.31	DNS		14.93	15.73	14.84	14.83								
140	DNS	DNS	14.65	DNS		15.03	16.19	15.1	14.83								
141	DNS	DNS	14.71	DNS		15.05	16.21	15.11	15.4								
142	DNS	DNS	14.83	DNS		15.11	16.28	15.34	15.61								
143	DNS	DNS	14.88	DNS		15.18	DNS	15.38	15.66								
144	DNS	DNS	14.92	DNS		15.34	DNS	15.5	15.68								
145	DNS	DNS	15.05	DNS		15.45	DNS	15.56	15.81								
146	DNS	DNS	15.12	DNS		15.77	DNS	15.63	16.13								
147	DNS	DNS	15.48	DNS		15.87	DNS	16.86	16.24								
148	DNS	DNS	15.52	DNS		16.05	DNS	DNS	16.44								
149	DNS	DNS	15.96	DNS		DNS	DNS	DNS	16.64								
150	DNS	DNS	16.51	DNS		DNS	DNS	DNS	17.32								
151	DNS	DNS	DNF	DNS		DNS	DNS	DNS	17.45								
152	DNS	DNS	DNS	NM		DNS	DNS	DNS	DNS								
153	DNS	DNS	DNS			DNS	DNS	DNS	DNS								
154	DNS	DNS	DNS			DNS	DNS	DNS	DNS								
155	DNS	DNS	DNS			DNS	NM	DNS	DNS								
156	DNS	NM	DNS			DNS	NM	DNS	DNS								
157	DNS		DNS			DNS	險俣	DNS	DNS								
158	DNS		DNS			DNS	險俣	DNS	DNS								
159	DNS		DNS			DNS		DNS	DNS								
160			DNS			NM		NM	險俣								
161			DNS					NM	險俣								
162								NM	險俣								
163									險俣								
164									險俣								
165									險俣								
166																	
167																	
168																	

③ 簡単に箱ひげ図を作ることができる

④ 次にここをクリックし 見やすい箱ひげ図をつくる

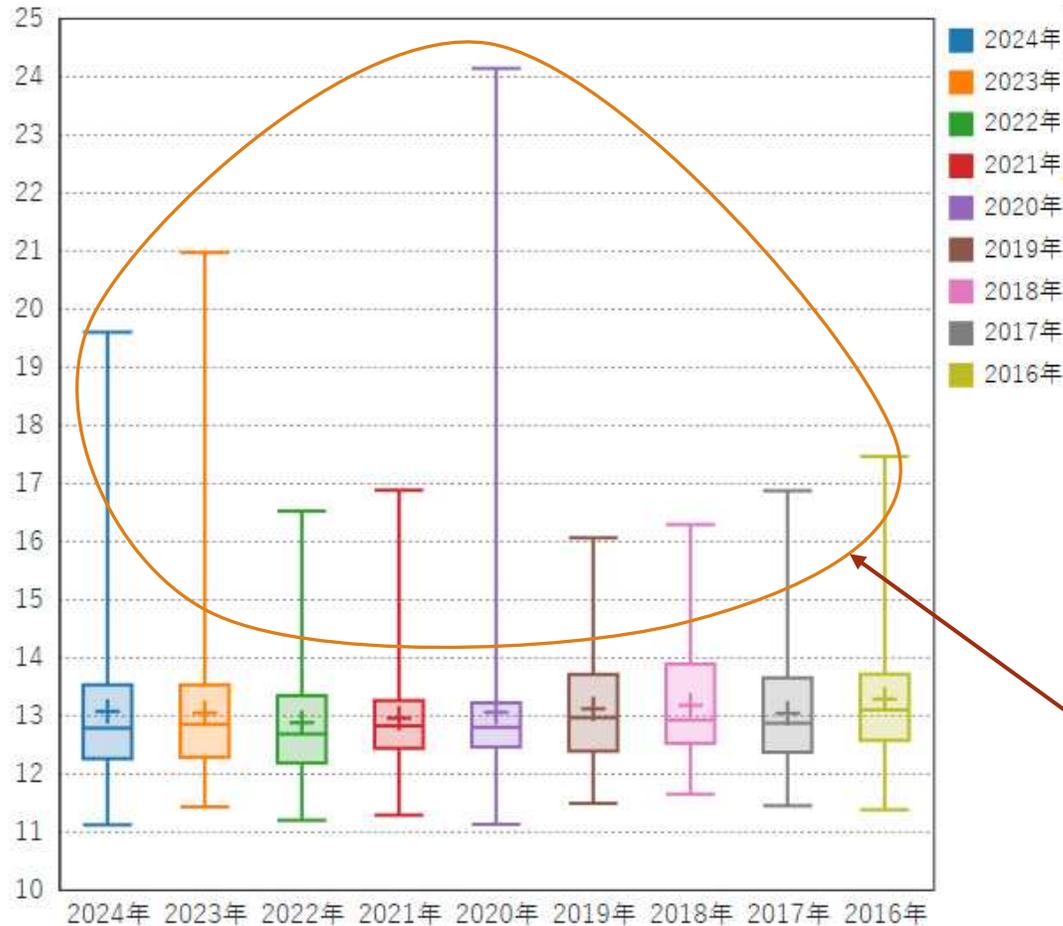
グラフ表示 100% 全画面表示

2024年 2023年 2022年 2021年 2020年 2019年 2018年 2017年 2016年

10.00 11.42 12.84 14.26 15.68 17.10 18.52 19.94 21.36 22.78 24.20

度数分布表 ヒストグラム 箱ひげ図

分析する(箱ひげ図)



階級 (秒)	2024年	2023年	2022年	2021年	2020年	2019年	2018年	2017年	2016年
	度数(人)								
平均値	13.063	13.038	12.868	12.947	13.048	13.108	13.168	13.031	13.276
中央値	12.775	12.84	12.675	12.815	12.79	12.955	12.915	12.86	13.09
最頻値(度数分布)	12.75	12.75	12.25	12.75	12.75	12.25	12.75	12.25	12.75, 13.25
最大値	19.59	20.96	16.51	16.87	24.13	16.05	16.28	16.86	17.45
最小値	11.11	11.42	11.19	11.28	11.12	11.48	11.64	11.44	11.37
第1四分位数	12.25	12.275	12.175	12.43	12.45	12.38	12.515	12.36	12.565
第3四分位数	13.52	13.52	13.335	13.25	13.21	13.7	13.88	13.64	13.705

上の表から

- 平均値で一番速いのは2022年
- 中央値で一番速いのも2022年
- 度数分布の最頻値で一番速いのも2022年
- 第1四分位数で一番速いのも2022年

ということで2022年が一番速いと言えそうだ。

でも、箱ひげ図を分析すると、全体のタイムと比較し、大きくかけ離れたタイムが出ている。

このタイムも入れて分析することが正しいのだろうか？

※大阪市春季総合体育大会では、チーム内で一番速い選手も出場するが、チーム内で決められた人数内であれば誰でも出場できるので、大きくかけ離れたタイムが出るときがある。

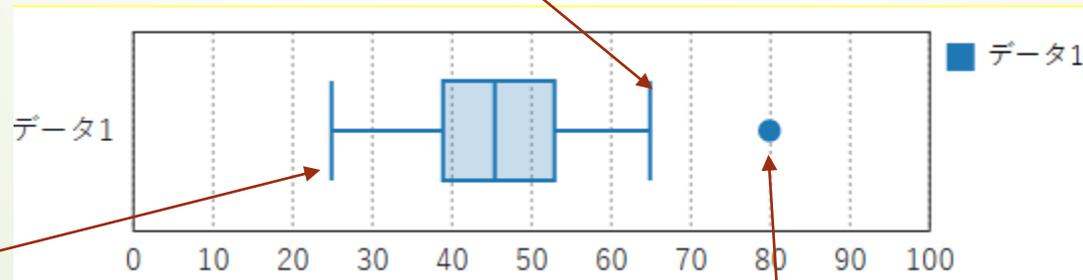
外れ値を見つけよう

他のデータと比べて極端な値のデータのことを、**外れ値**と呼びます。
外れ値を含んだ状態のままデータ分析を行うと、外れ値に引っ張られて平均値などの分析結果や傾向が変わってしまう場合があるため、取り扱いには注意が必要です。

箱ひげ図を使った外れ値の見つけ方

四分位範囲 = 第3四分位数(Q3)-第1四分位数(Q1)

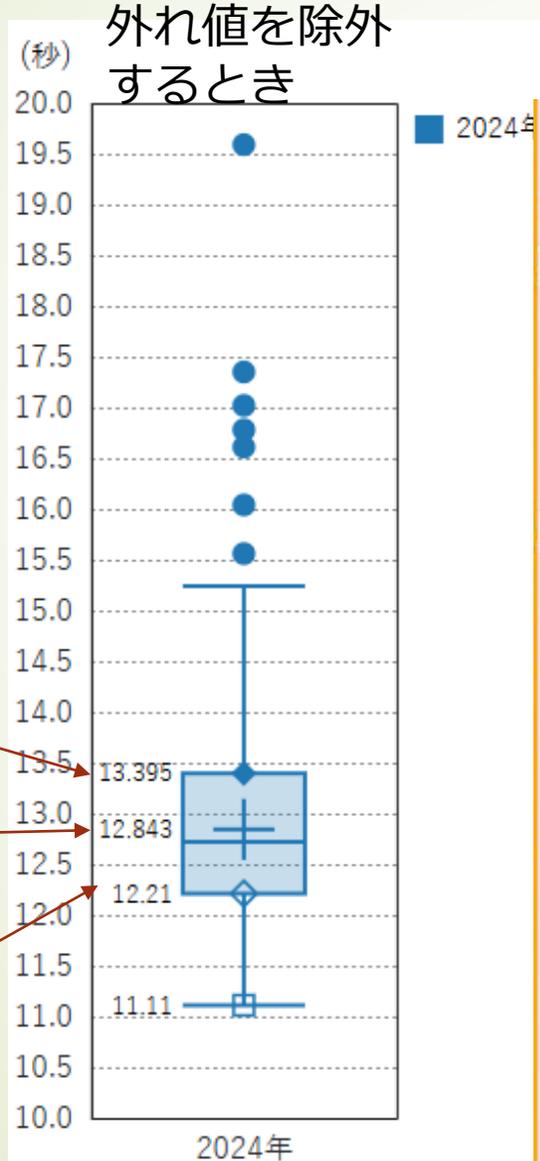
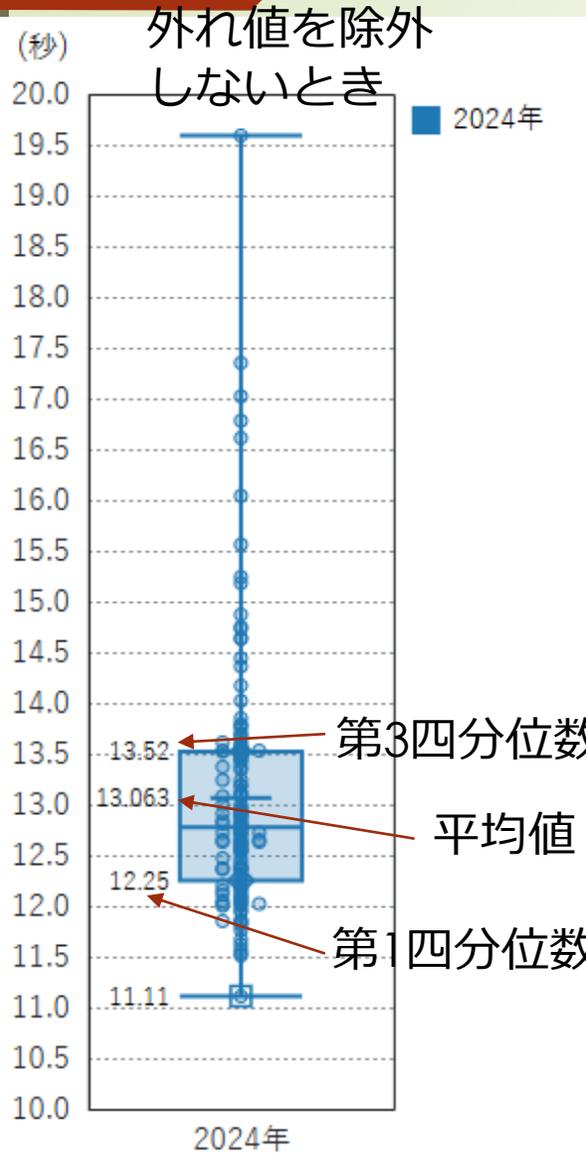
ひげの右端 = 第3四分位数+四分位範囲×1.5 より小さい最大値



ひげの左端 = 第1四分位数-四分位範囲×1.5 より大きい最小値

ひげの左端・右端に収まらない値を**外れ値**とします

SGRAPAで箱ひげ図の外れ値を除外しよう



縦向きにする

凡例を表示する 大きさ 14

グラフの色

セルの範囲	データ名 (凡例)	色
A1:A134	2024年	

すべてのグラフを同色にする

表示する項目

すべて選択

<input checked="" type="checkbox"/> 平均値	(<input checked="" type="checkbox"/> 値 <input checked="" type="checkbox"/> マーカー)
<input type="checkbox"/> 最小値	(<input type="checkbox"/> 値 <input type="checkbox"/> マーカー)
<input checked="" type="checkbox"/> 第1四分位数	(<input checked="" type="checkbox"/> 値 <input checked="" type="checkbox"/> マーカー)
<input type="checkbox"/> 第2四分位数	(<input type="checkbox"/> 値 <input type="checkbox"/> マーカー)
<input checked="" type="checkbox"/> 第3四分位数	(<input checked="" type="checkbox"/> 値 <input checked="" type="checkbox"/> マーカー)
<input type="checkbox"/> 最大値	(<input type="checkbox"/> 値 <input type="checkbox"/> マーカー)
<input type="checkbox"/> ドットプロット	

外れ値を データから除外しない データから除外する

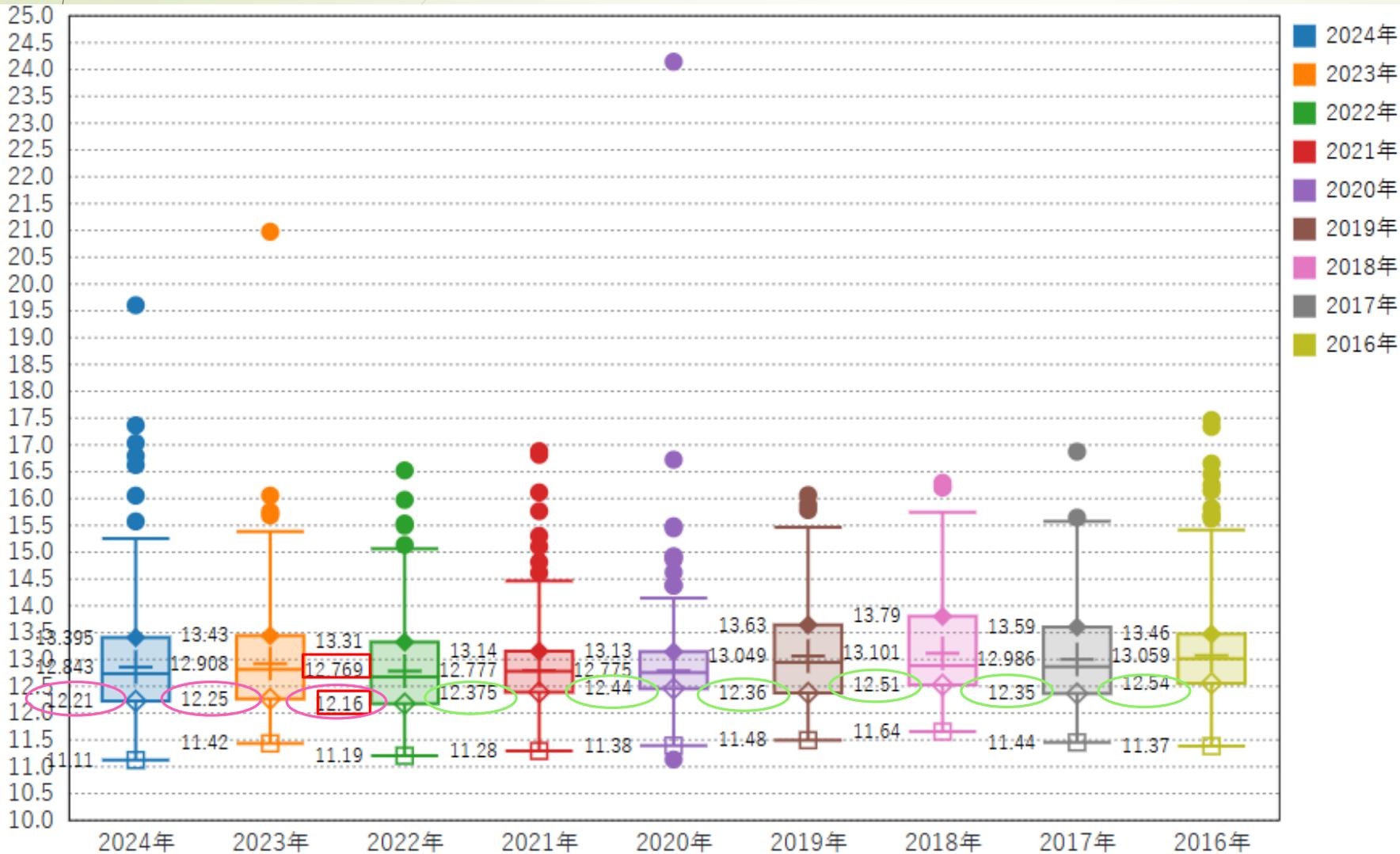
外れ値の設定 ユーザが設定した値 自動で判定した値

SGRAPAでは外れ値を簡単に除外する機能がある

まったく同じデータでも外れ値を除外するかしないかで、平均値や第1四分位数等が変わってくる。

外れ値を除外したデータで比較してみよう

外れ値を除外して分析する(箱ひげ図)



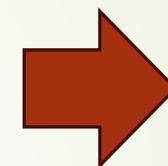
外れ値を除外して分析してみると

- 平均値で一番速いのは2022年
- 第1四分位数(上位1/4)が速いのも2022年
- 2021年以前と2022年以降を比較すると第1四分位数(上位1/4)が2022年以降の方が速い
 - クラブチームや陸上動画の影響?
- 2020年はコロナの影響があったのに例年並みに善戦
 - 8月中旬開催だった
 - 部活動再開後2か月程度だったのに選手はよく頑張った

このデータ分析は正確な分析といえるか？

→ 風の影響

- 1.0mの追風が吹くと0.1秒タイムが速くなるといわれている
- 実際2018年は向かい風のレースが多かった
- 2016年は強い追い風の中でレースが行われた
- 風の影響を考えない補正タイムで比較する必要がありそうだ



風の影響を受けないタイムを計算するサイトもあるが、今回は膨大なデータのため計算をしていない

(野口純正 / 『100m 9秒台・200m 19秒台の過去・現在・近未来』 / 「フューチャーアスレティックス・第2号」 / 近未来陸上競技研究所編 / 2002年12月に掲載)

2024		2023		2022		2021		2020		2019		2018		2017		2016	
記録	風	記録	詳細	記録	詳細	記録	詳細										
11.11	1.2	11.42	-0.4	11.19	0.6	11.28	2.4	11.12	0.4	11.48	-0.6	11.64	2.5	11.44	2.2	11.37	2
11.51	0.9	11.47	-0.1	11.59	0.2	11.54	0.3	11.38	0.4	11.51	0.1	11.67	-1.3	11.47	2.8	11.47	4.4
11.53	0.3	11.6	0.3	11.66	1.5	11.57	-0.6	11.55	0.4	11.55	-0.6	11.71	-1.8	11.5	1.6	11.49	1.7
11.57	3.1	11.64	-0.4	11.67	0.5	11.58	0.8	11.77	1.2	11.72	0.1	11.86	0	11.54	-0.2	11.62	4.3
11.62	1.4	11.66	-0.1	11.67	0.1	11.63	0.3	11.82	0.4	11.81	0.3	11.9	1.2	11.6	1.2	11.71	1.6
11.67	0.5	11.7	-0.1	11.72	1.2	11.69	1.5	11.86	-1.2	11.82	0.6	11.91	1.5	11.67	0	11.75	0.6
11.76	-1.5	11.73	0.2	11.74	-0.2	11.71	-0.8	11.93	-1.2	11.82	-0.6	11.92	-1.8	11.73	3.4	11.77	3.5
11.81	-0.4	11.74	0.3	11.77	0.9	11.77	0.7	11.96	0.6	11.82	-0.5	11.93	-0.9	11.75	0.3	11.8	2.3
11.83	0.9	11.75	-0.4	11.81	0.8	11.81	-0.3	11.98	0.4	11.85	0	11.95	-2.2	11.75	0.7	11.92	1.7
11.85	0	11.76	-0.8	11.83	1	11.82	2.1	12.04	0.4	11.87	-1.4	11.95	-0.2	11.8	-1.3	12	3.5
11.85	-0.3	11.78	0.5	11.86	0.8	11.83	0	12.05	0.4	11.93	-0.6	11.98	-2.5	11.82	0.8	12.01	3.5
11.86	0.9	11.85	0.7	11.88	0.4	11.9	-0.3	12.09	0.4	11.94	-0.7	11.98	-1.6	11.85	3.1	12.02	1.8
11.94	-0.3	11.88	0.2	11.88	0.3	11.92	-0.6	12.12	0.6	11.95	-1.2	12.04	-2.7	11.89	2.7	12.08	1.5
11.96	-2.2	11.93	0.3	11.88	0.8	11.93	1.2	12.15	-1.2	11.97	-1.2	12.05	0.7	11.9	3.1	12.12	4.3
12	1.4	11.94	0.3	11.89	1.2	11.97	-0.3	12.16	-1.2	12.01	0	12.08	-0.2	11.91	2.7	12.12	2.9
12	0.3	11.95	0.7	11.92	-0.1	11.98	1.6	12.16	-1.7	12.04	-0.4	12.1	-1.6	11.94	1.5	12.16	1.9
12.02	0.8	11.97	0.5	11.95	0.6	12	-0.6	12.18	-1.7	12.09	-1.1	12.17	1.2	11.95	-0.2	12.17	3.5
12.02	-2.2	12.01	0.2	11.96	-0.1	12.01	0.2	12.21	-1.2	12.12	-0.5	12.21	-3	11.99	0.3	12.18	4.4

部活動 × 統計 = 甲子園？

滋賀県立膳所高校野球部

2018年 選抜高校野球大会に出場

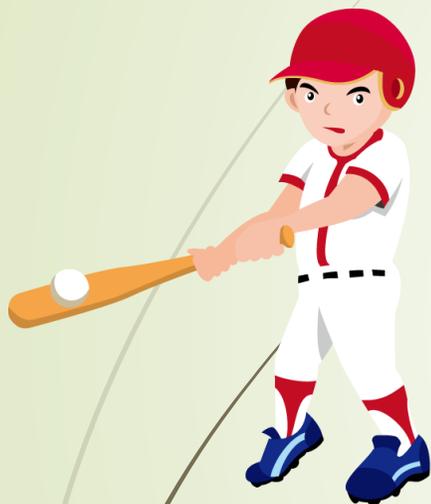
「選手のプレーの技量的にはうまくない。その分、工夫して練習や試合に役立てたい」と、野球データを統計学的見地から分析する「セイバーメトリクス」を取り入れ、選手起用や試合戦術に生かすように。

秘密兵器は「データ班」部員

試合に臨む前に、対戦相手校のデータを取り、客観的な数値で特徴を洗い出す。各打者がどの球種をどの方向へ、どんな打球を飛ばしたのか、結果を詳細にまとめ、傾向を分析し、選手に提供。

「中堅へ抜けていきそうないい打球を打ち返されても、振り返ると野手が二塁ベースの後ろにいて取ってくれたり、外野手が守備位置を極端に寄ってくれていたおかげで普通なら長打コースの打球を単打で止められたりした。思いがけない形で助けてもらった。」と効果を実感。

全国屈指の進学校が甲子園切符を勝ち取った理由とは 将来の夢「普通のお嫁さん」の女性部員が秘密兵器
<https://www.sankei.com/article/20180208-B2SGZ6U75FKEBELLU603POV4KY/>



部活動 × 統計 = 日本中学新記録更新へ

米子高専 短距離走支援アプリ 開発



スマートフォンで撮影した映像から、
いろんな関節の角度を計算したりと
か、伸展角度を計算したりだとかして、
ユーザーに自分の走りがどんな
特性を持っているのかというのを提
供できるサービスを目指しています。

米子高専 矢田ほのかさん



広島県中学3年生 三好美羽さん

👉今年度100m走で日本中学新記録

「お～、え～なんかすごい、面白
いです。今後直すところとか、設
置の角度とかはできるだけフラッ
トにいった方がいいので、そうい
うところで分析していただけると、
本当に助かります」

「俊足女子中学生・三好美羽選手のフォームの秘密は？ 米子高専生が映像解析し、短距離走支援アプリ開発に挑む」より
<https://newsdig.tbs.co.jp/articles/bss/1166907?display=1>

このように統計を身近なところに活用していくと、より良い生活につながっていくことを子どもたちに実感させて
いきたい。

参考資料等

- SGRAPA

URL: <https://sgrapa.com/>

(こちらのソフトを使ってグラフなどを作りました)

- 大阪市中体連陸上競技部

URL: <https://citytyutairen.sakura.ne.jp/>

(リザルトのページから各年の春季総合体育大会の男子100mのデータを取り出しました)

- 産経新聞

全国屈指の進学校が甲子園切符を勝ち取った理由とは 将来の夢「普通のお嫁さん」の女性部員が秘密兵器

URL: <https://www.sankei.com/article/20180208-B2SGZ6U75FKEBELLU6O3PQV4KY/>

- 山陰放送ホームページ

俊足女子中学生・三好美羽選手のフォームの秘密は？ 米子高専生が映像解析し、
短距離走支援アプリ開発に挑む

URL: <https://newsdig.tbs.co.jp/articles/bss/1166907?display=1>

- イラスト：イラストポップ

URL: <https://illpop.com/>

- 東洋経済ONLINE キャリア・教育 確かなリアルが見えてくる「統計」超入門

データ分析始めると必ず悩む「外れ値」の壁克服法 集計したデータはそのまま使ってはいけない

URL: <https://toyokeizai.net/articles/-/539717>