

2 データの整理

テーマ「2023年の福崎町は本当に暑かったのか？」

[福崎町で 38.9 度 20 日の国内最高気温に同町の観測史上最高を更新県内 21 日以降も真夏日が続く見込み | 社会 | 神戸新聞 NEXT \(kobe-np.co.jp\)](https://www.kobe-np.co.jp/news/society/202308/0016721209.shtml)

<https://www.kobe-np.co.jp/news/society/202308/0016721209.shtml>

2023年8月20日に、福崎町は日本で最高気温となる38.9℃を記録しました。この年の夏のニュースでは、エルニーニョ現象が伝えられました。しかし、2023年の福崎町は本当に暑かったのでしょうか？2022年と2023年の7月と8月の日最高気温（1日における最高気温）の傾向の違いを統計的に分析してみましょう。

学習活動

福崎町の2022年と2023年の7月と8月の日最高気温のデータを気象庁からダウンロードし、一つの表（データ）にまとめた。このデータの外れ値を求める。

収集したデータをそのまま分析ができるわけではない。データの中には、数値が欠けている欠損値や他の値から大きく外れた外れ値が含まれているかもしれない。分析の行う前に、集めたデータを確認しておくことが重要である。

外れ値

今回は、「第1四分位数-1.5×四分位範囲」より小さい値、「第3四分位数+1.5×四分位範囲」より大きい値を「外れ値」として扱う。

※四分位範囲＝第3四分位数－第1四分位数

外れ値の求め方

（方法1）関数から外れ値を求める

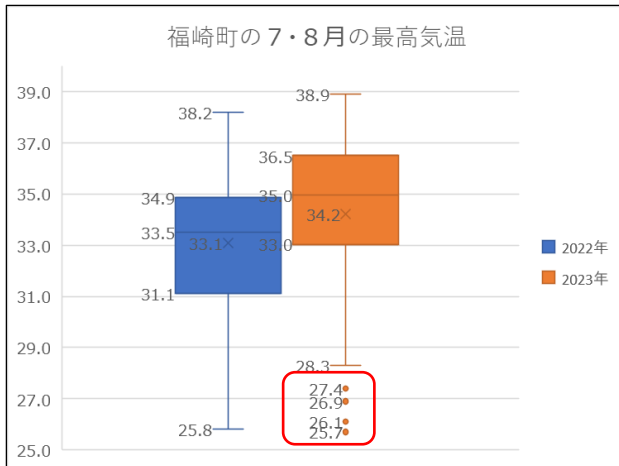
	2022年	2023年
第3四分位数+1.5×四分位範囲	40.5	41.7
第3四分位数	34.9	36.5
中央値	33.5	35.0
第1四分位数	31.1	33.0
第1四分位数-1.5×四分位範囲	25.4	27.8

中央値（第2四分位数）、第1四分位数、第3四分位数を求めるための数式例

中央値
=median(範囲) または =QUARTILE.INC(配列, 2)
第1四分位数
=QUARTILE.INC(配列, 1)
第3四分位数
=QUARTILE.INC(配列, 3)

※範囲、配列は、外れ値を求めるための該当データ

(方法2) 箱ひげ図から外れ値を求める



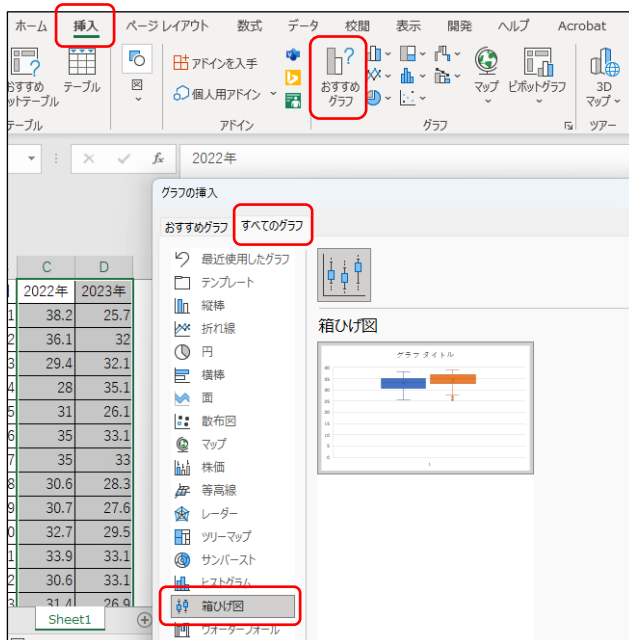
表計算ソフト (Excel 2019) で箱ひげ図を作成する

1. データを選択する

B	C	D
日	2022年	2023年
1	38.2	25.7
2	36.1	32
3	29.4	32.1
4	28	35.1
5	31	26.1
6	35	33.1
7	35	33
8	30.6	28.3
9	30.7	27.6
10	32.7	29.5
11	33.9	33.1
12	30.6	33.1
13	31.4	26.9

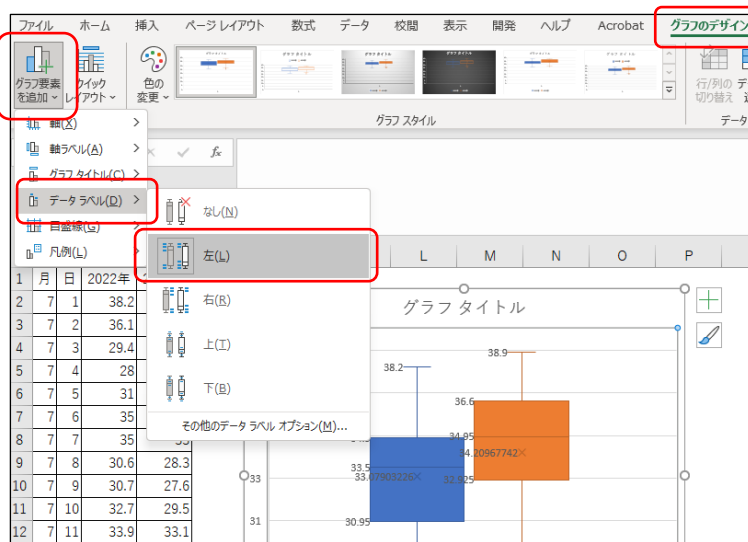
2. グラフを挿入する

([挿入]-[おすすめグラフ]-[すべてのグラフ]-[箱ひげ図])



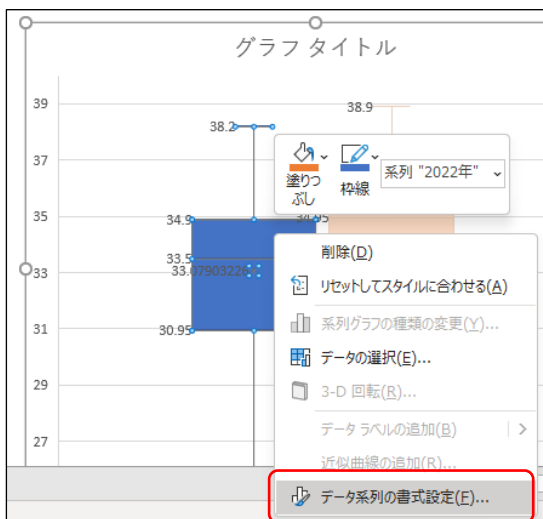
3. データラベルを表示する

([グラフデザイン]-[グラフ要素を追加]-[データラベル])

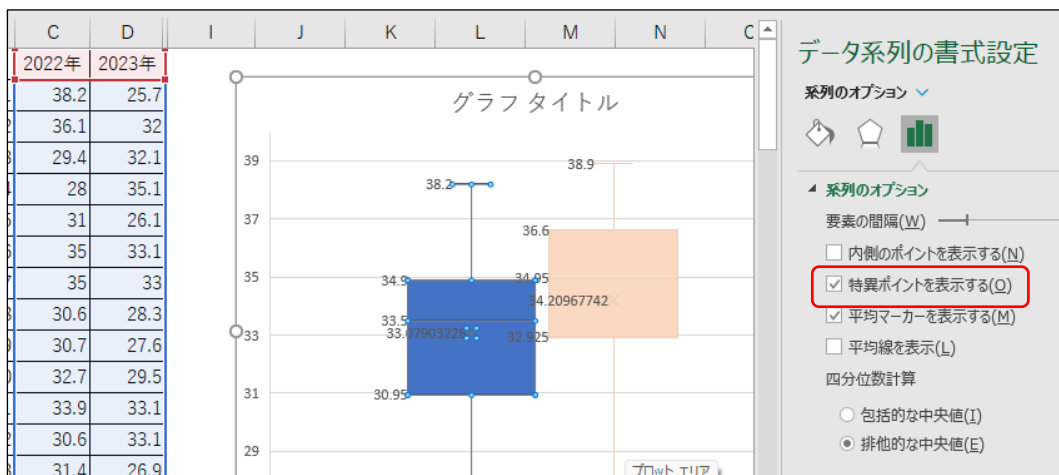


4. 外れ値 (特異ポイント) を表示する

(該当グラフを右クリック-[データ系列の書式設定])



「特異ポイントを表示する」にチェックを入れる



(留意点)

外れ値を除いて統計処理する機会が多いが、外れ値が入力ミス等による異常値か、ある条件が重なったときに現れる値かは、分析結果を比較した上で判断する必要がある。外れ値を求める他の方法として、クラスター分析などがある。

学習評価例

項目	ねらい	重点	備考
1	○データの整理 データに含まれる外れ値の扱いやデータを整理する必要性を理解し、適切にデータを扱うことができる。	知	

重点…重点的に生徒の学習状況を見取る観点を示している。