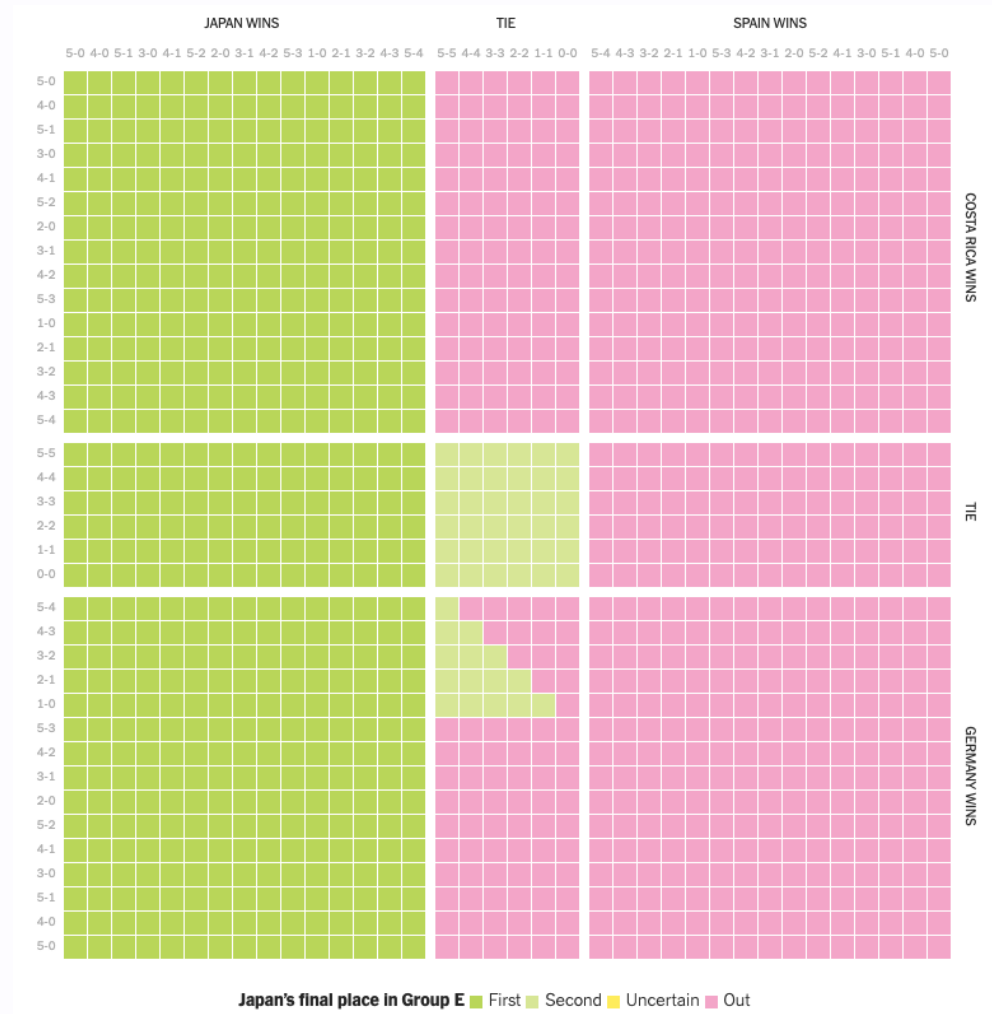


Introduction to Statistics

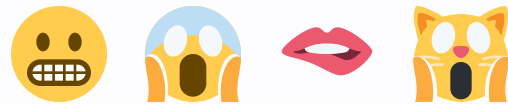
統計学入門

Week 10 | December 1, 2022



<https://www.nytimes.com/interactive/2022/upshot/japan-world-cup-scores-standings.html>

Week 9 小テスト



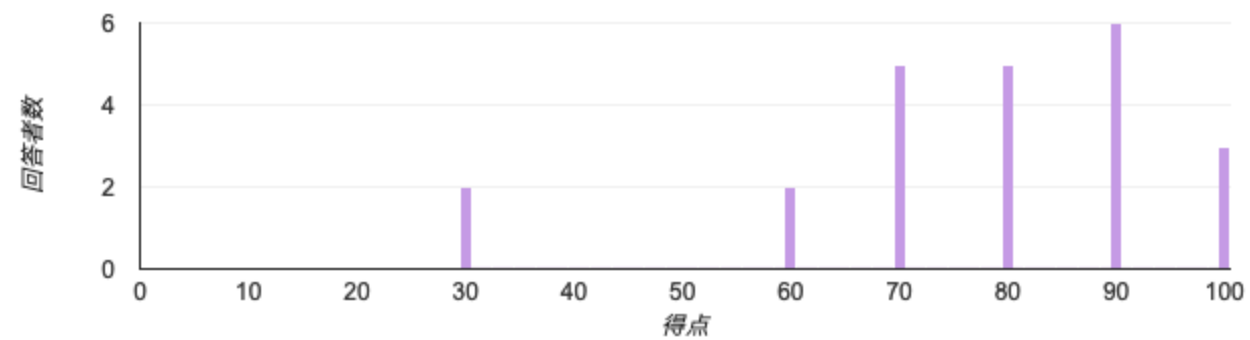
分析情報

平均
76.96/100 ポイント

中央値
80/100 ポイント

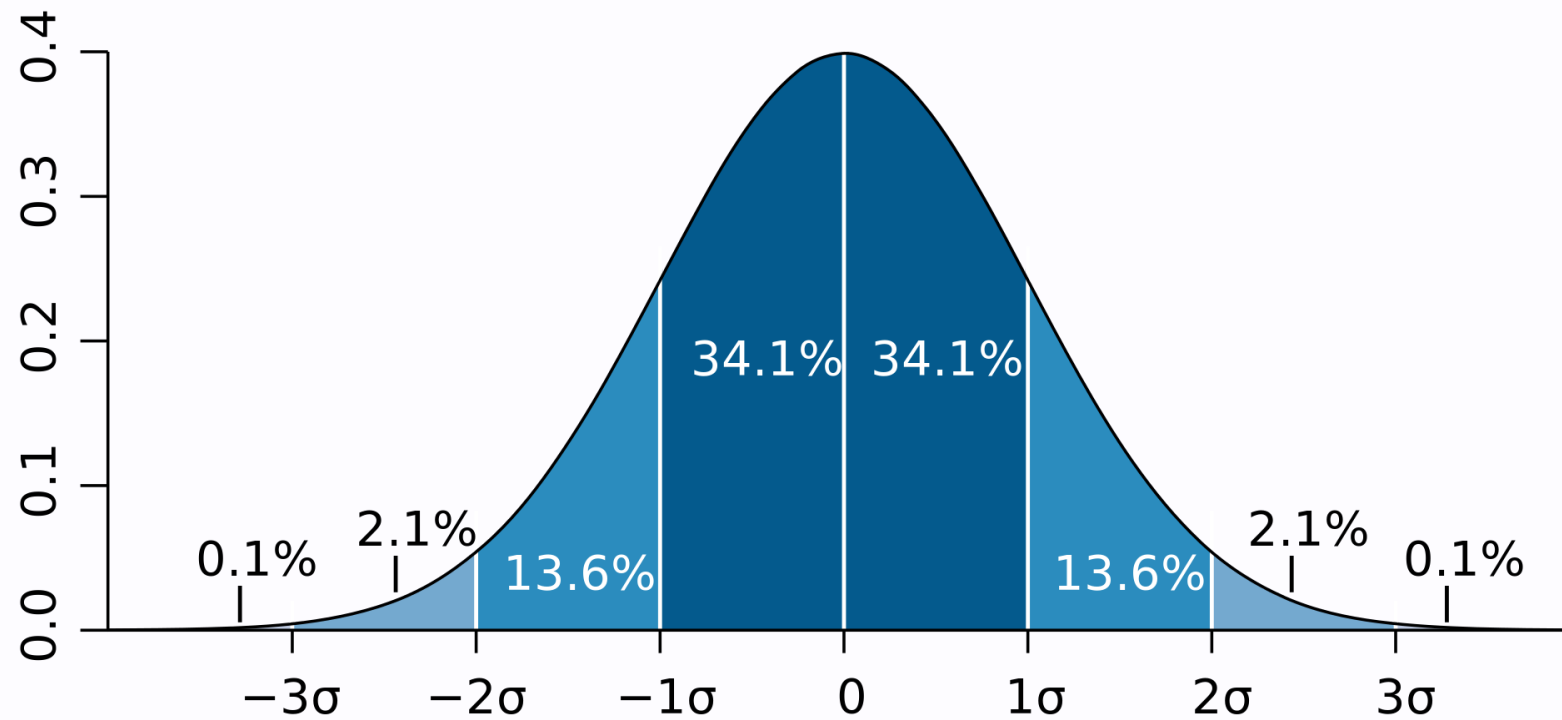
範囲
30～100 ポイント

合計点の分布



Last week

正規分布

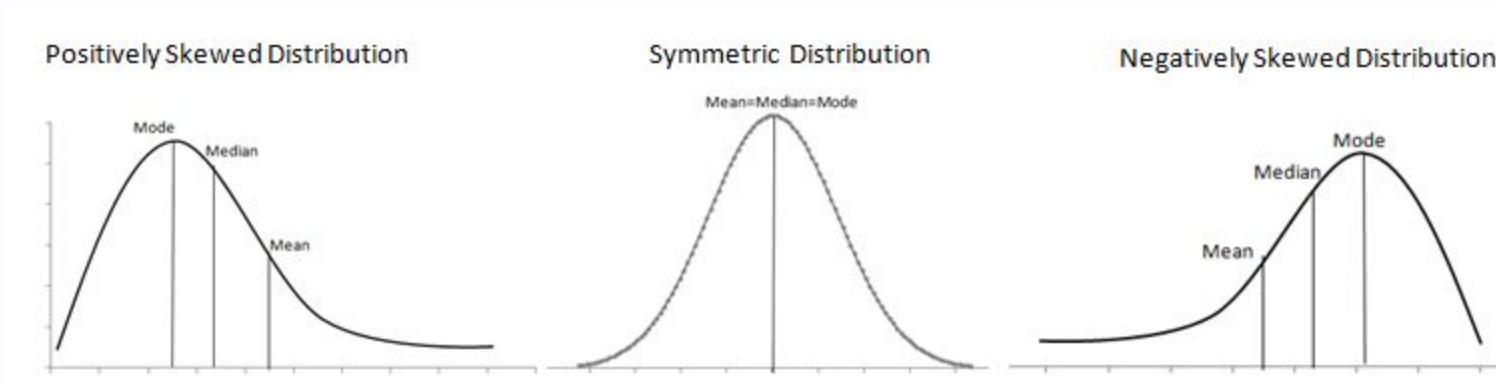


The "bell" curve

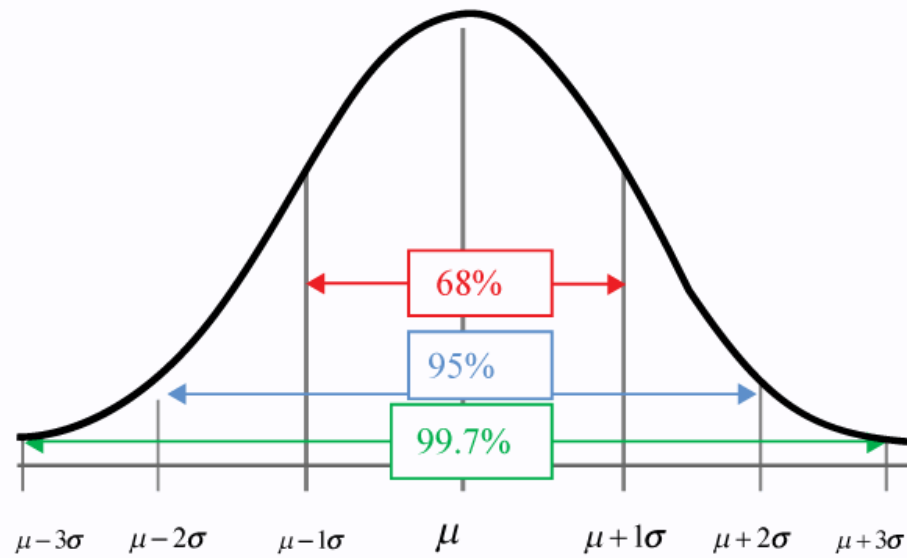
チェビシェフの不等式

Chebyshev's inequality

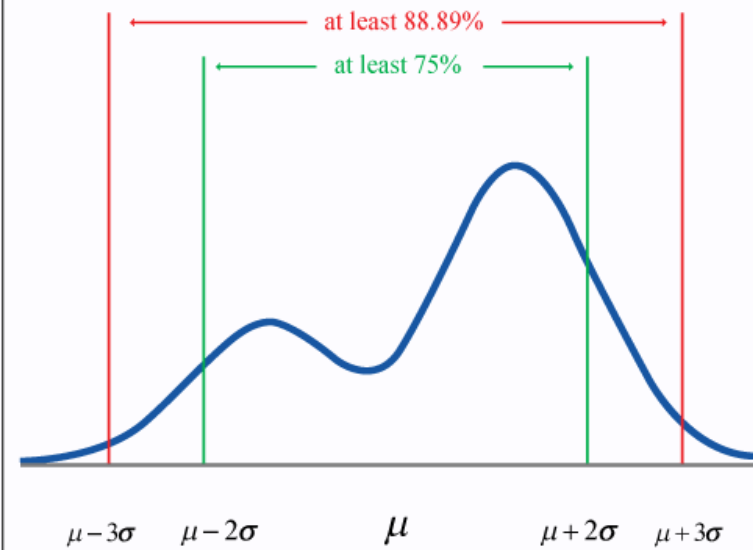
What if data was not normally distributed?



Empirical Rule (Normal Distributions)



Chebyshev's Inequality (Any Distribution)



変動係数

Coefficient of Variation

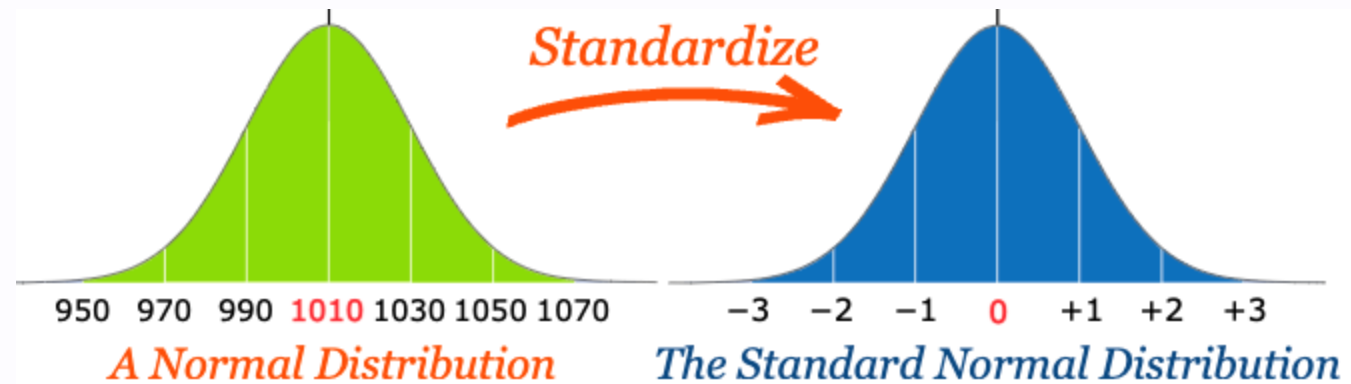
分布の広がりを表現する統計量

$$CV = \sigma / \mu$$

變動係數 = 標準偏差 / 平均值

標準化とZ値

Z-Score



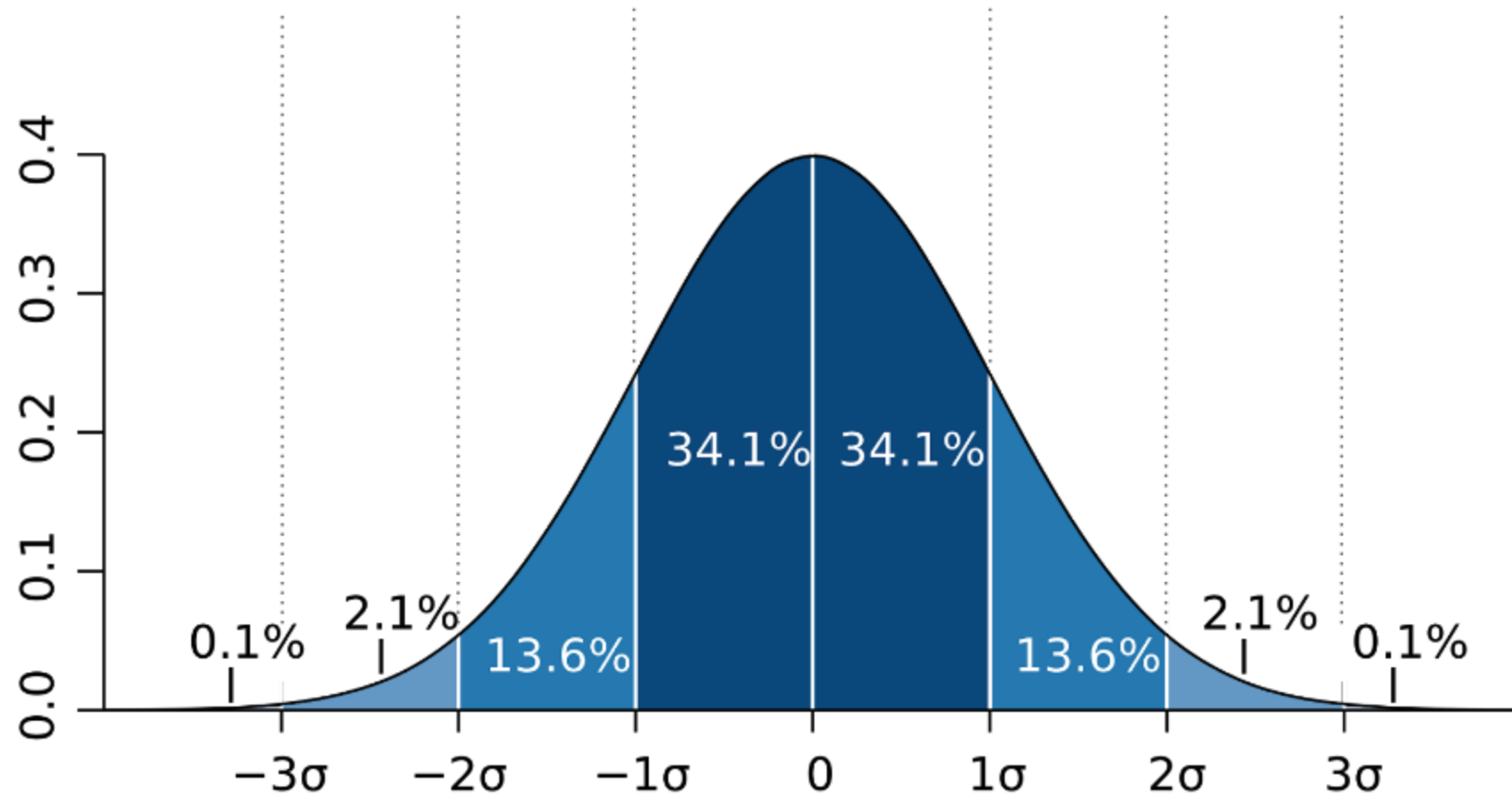
このクラスの平均身長

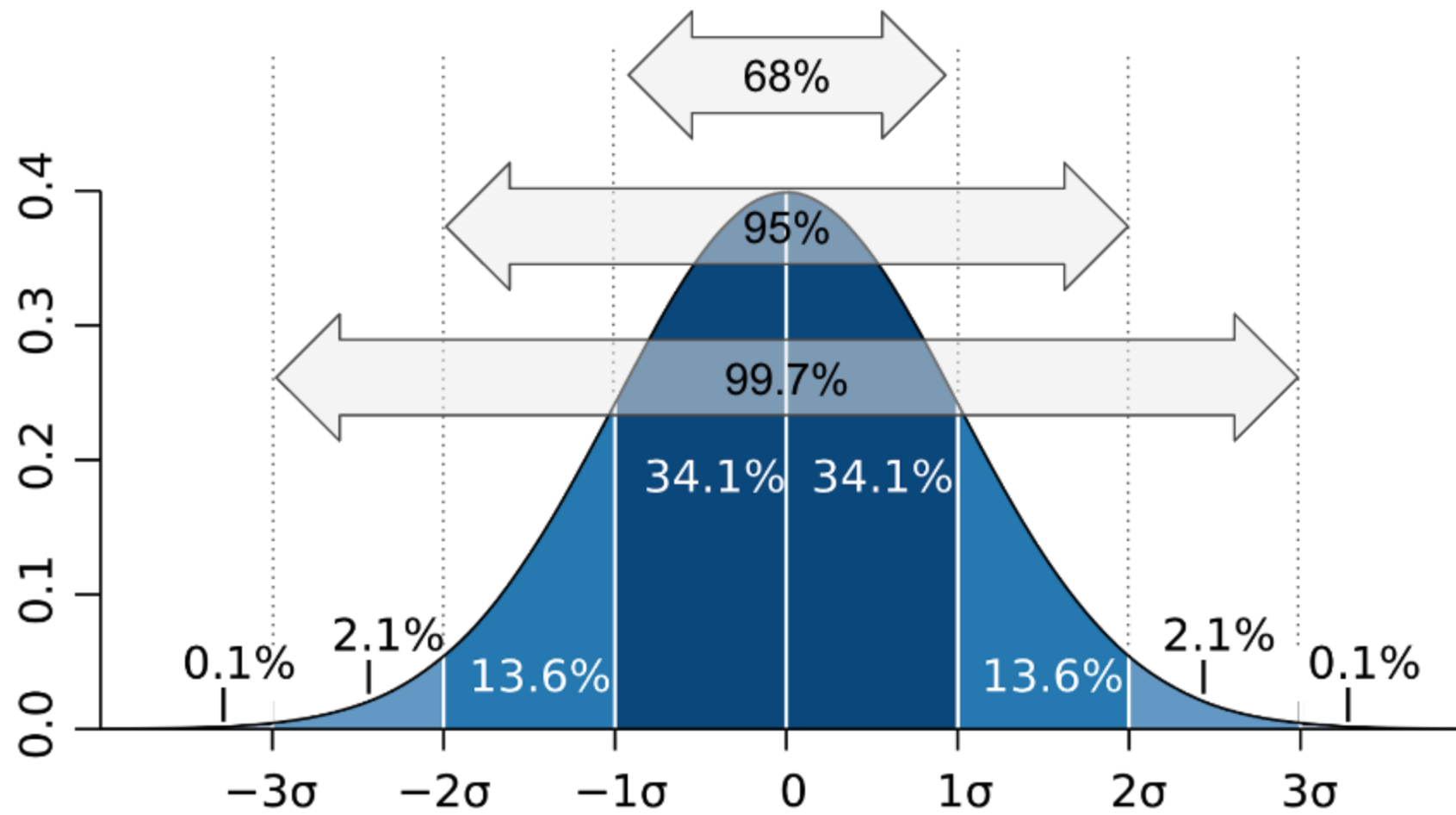
165cm(μ)

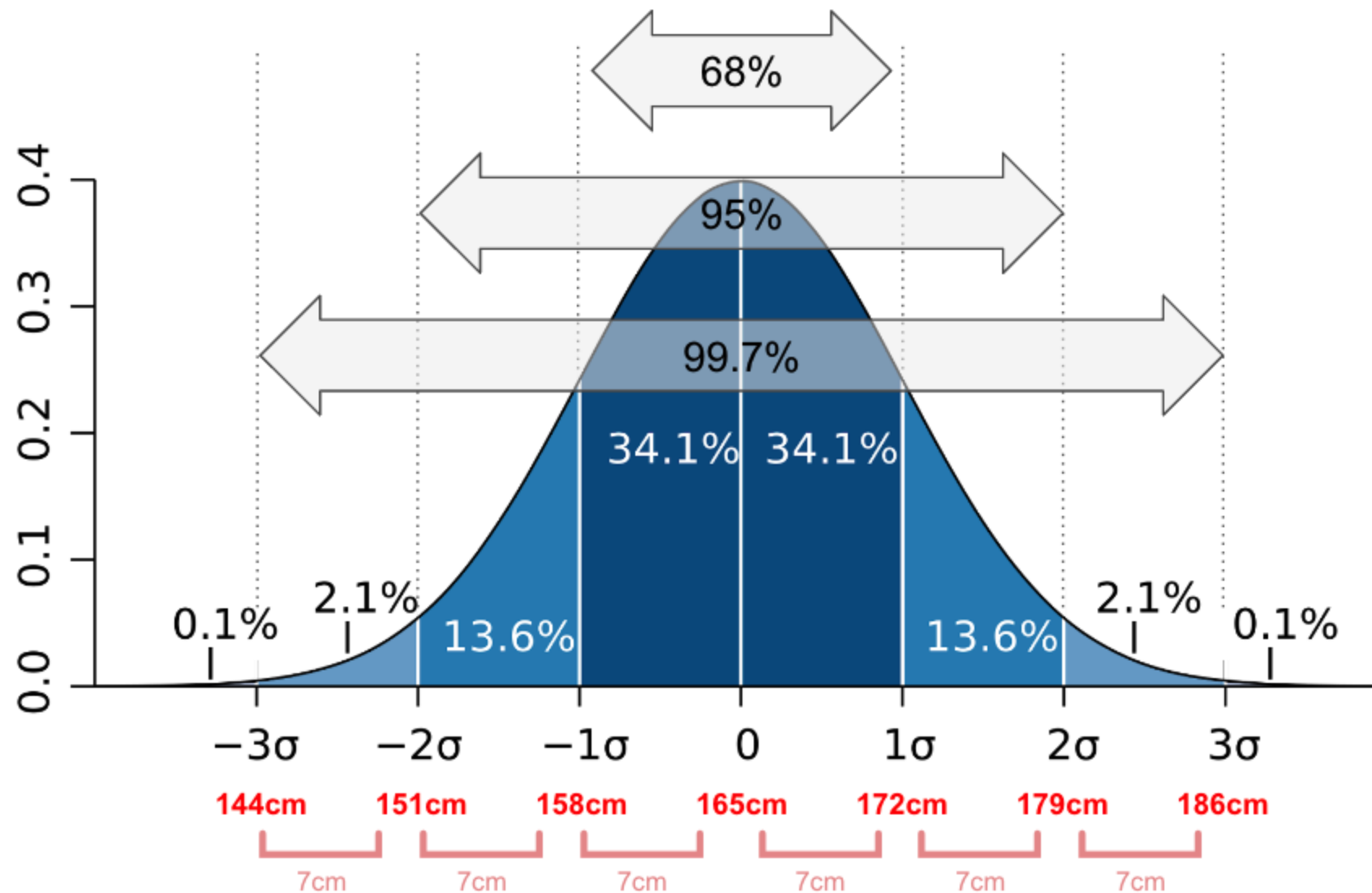
標準偏差

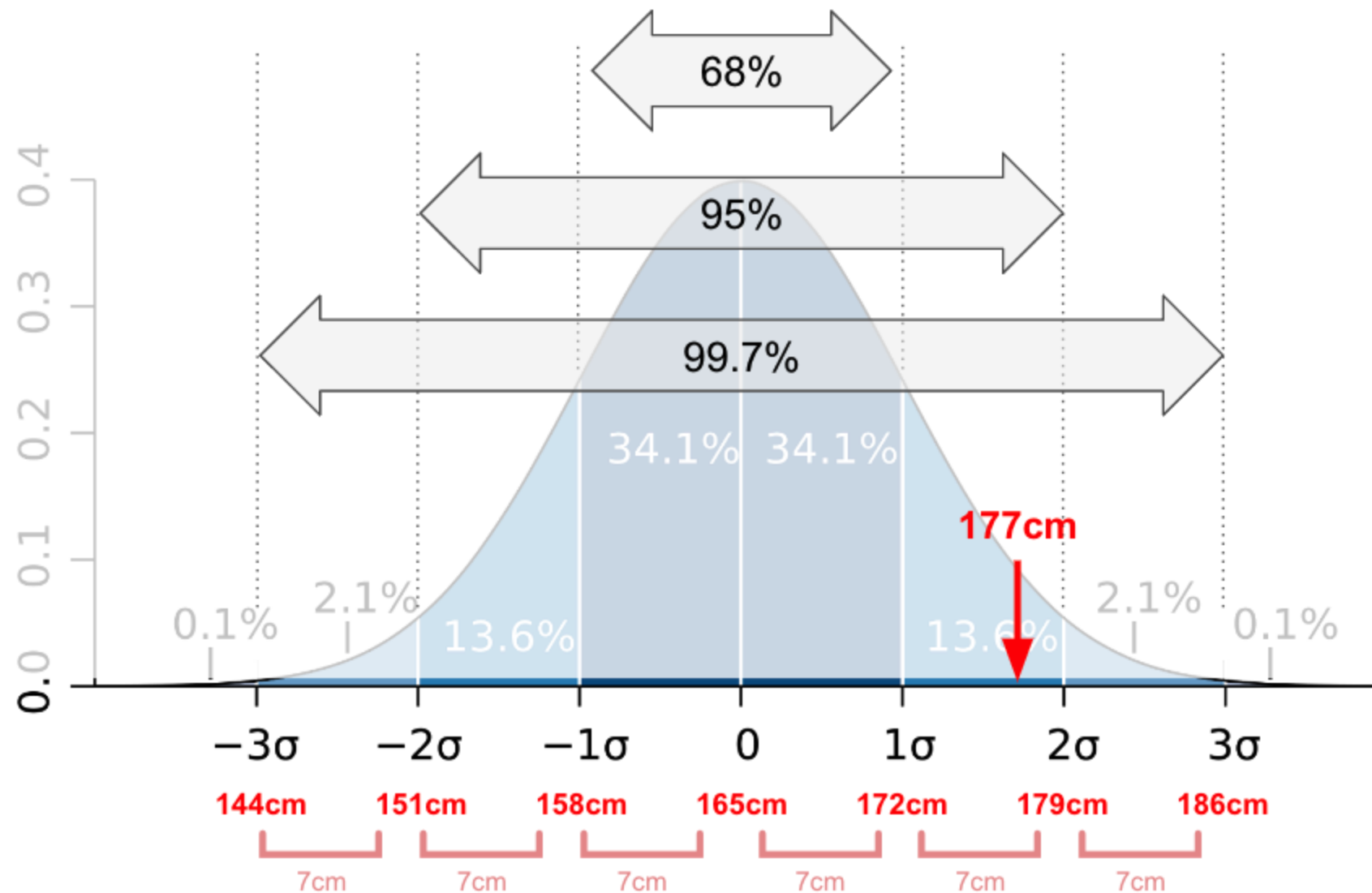
7cm(σ)

先生の身長は177cm ➡ 標準化すると？





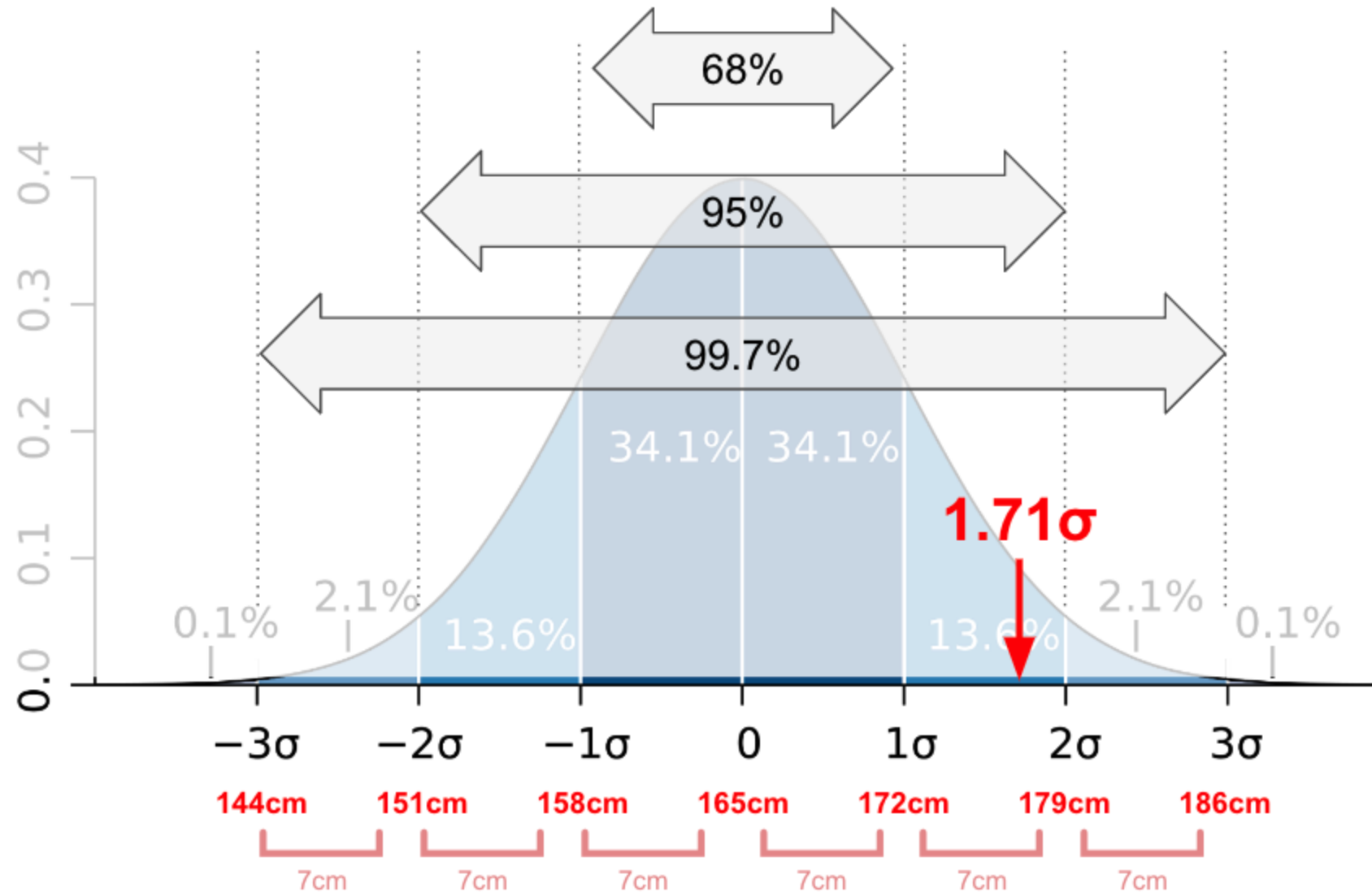




$$(177-165)/7 = 1.71$$

(求める身長-平均値)/標準偏差 = 標準化した変数

すなわち



散らばりのグラフ表現

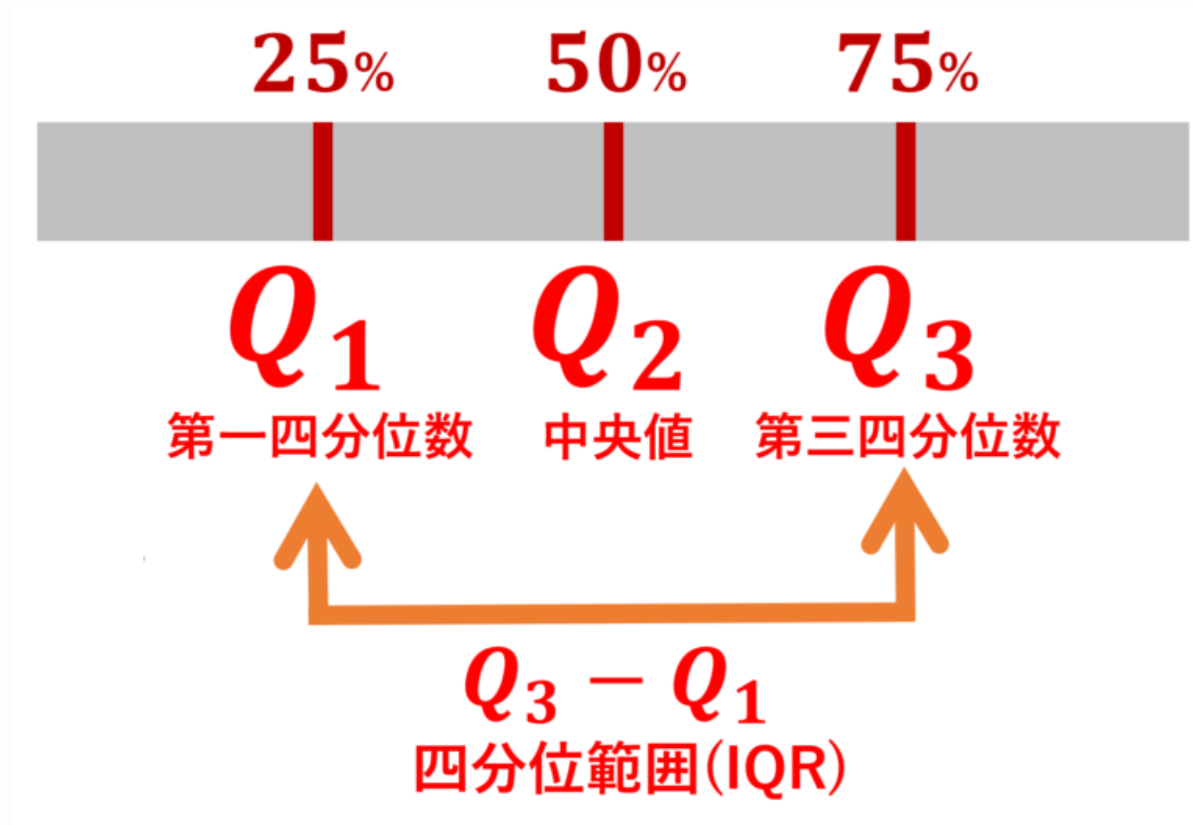


Image source: <https://mathlandscape.com/box-plot/>

データの5数要約を計算する：

四分位 = quartile

0. 最小値 (MIN)
1. 第1四分位点 (Q1)
2. 中央値 (Q2)
3. 第3四分位点 (Q3)
4. 最大値 (MAX)

Boxをかく

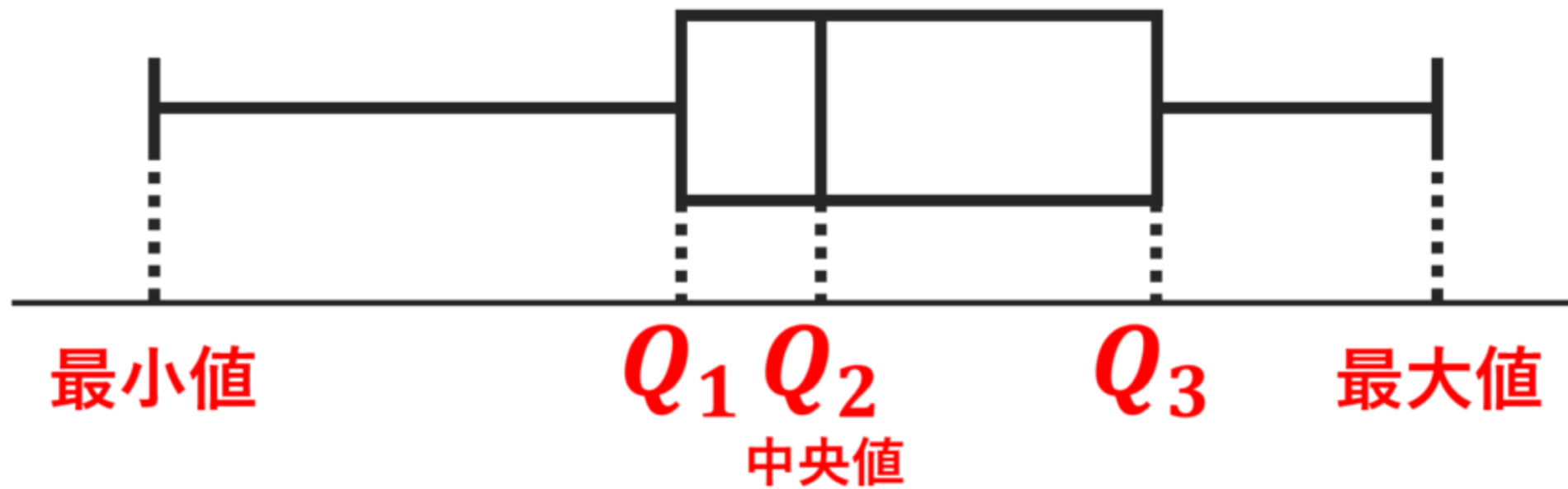


Image source: <https://mathlandscape.com/box-plot/>

ではやってみよう

名前	性別	身長 (cm)
Ran Komaba	Male	165
Makoto Kishino	Male	165
Jo Yamaguchi	Male	165
Kensuke Wakabayashi	Male	167
Sota Kobori	Male	170
Kuuta Uchida	Male	171
Yusuke Aono	Male	174
Tomoki Wada	Male	175
Masakazu Tamagawa	Male	175
Towa Tanahashi	Male	178

Box with whiskers (外れ値)

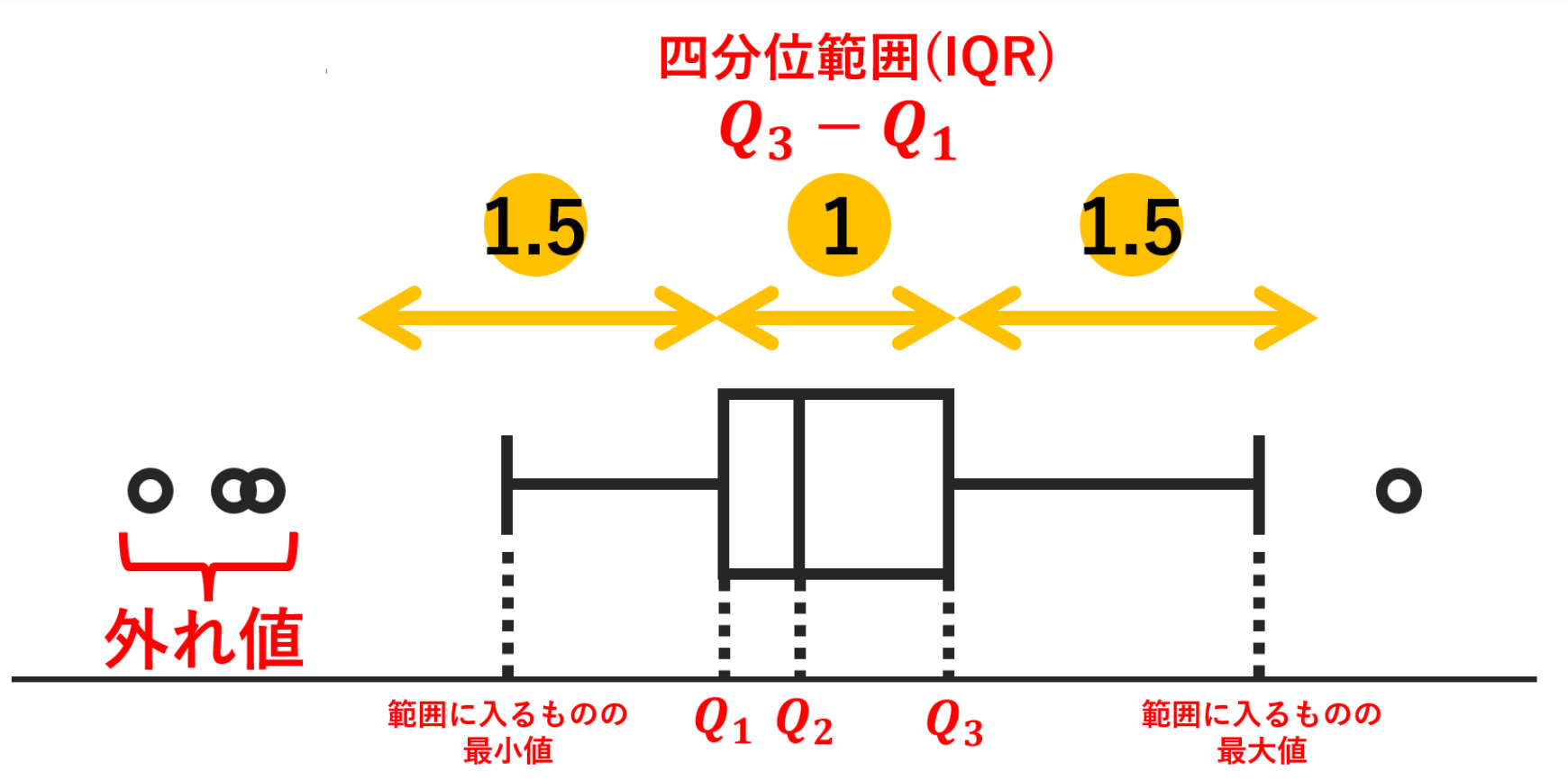


Image source: <https://mathlandscape.com/box-plot/>

Gol D. Roger
285 cm



Average man
In France (179 cm)
Or almost 6 feet

Zephyr
285 cm



Garp
290 cm



Animeの身長データで やってみよう

最小値 (MIN)	8
第 1 四分位置 (Q1)	108
中央値 (Q2)	153
第 3 四分位置 (Q3)	175
最大値 (MAX)	290

キャラクター名	身長 (cm)
ハム太郎	8
目玉のオヤジ	10
アリエッティ	10
ピカチュウ	40
ケロロ軍曹	55
ヨーダ	66
たらちゃん	95
ドラミちゃん	100
コナン	102
野原慎之介	106
アンパンマン	110
孫悟空	123
ドラえもん	129
鬼太郎	130
磯野ワカメ	133
鉄腕アトム	135
碇シンジ	141
磯野カツオ	143
セーラームーン	152
クリリン	153
初音ミク	158
フグ田サザエ	159
ベジータ	164
ガチャピン	165
峰不二子	167
アムロレイ	168
黒子テツヤ	168
フグ田マスオ	173
ルフィ	175
孫悟空	175
磯野波平	178
ルパン三世	179
ゴルゴ13	182
キン肉マン	185
桜木花道	189
ラーメンマン	209
ビッコロ	224
Zephyr	285
Garp	290

最小值 (MIN)	8
第 1 四分位置 (Q 1)	108
中央值 (Q 2)	153
第 3 四分位置 (Q3)	175
最大值 (MAX)	290

最小值 (MIN)	8
第 1 四分位置 (Q 1)	108
中央值 (Q 2)	153
第 3 四分位置 (Q3)	175
最大值 (MAX)	290

$$\begin{array}{c}
 \boxed{175} \\
 \hline
 \text{Q3}
 \end{array}
 -
 \begin{array}{c}
 \boxed{108} \\
 \hline
 \text{Q1}
 \end{array}
 =
 \begin{array}{c}
 \boxed{67} \\
 \hline
 \text{四分位範圍}
 \end{array}$$

最小值 (MIN)	8
第 1 四分位置 (Q 1)	108
中央值 (Q 2)	153
第 3 四分位置 (Q3)	175
最大值 (MAX)	290

$$\begin{array}{c} \text{175} \\ \hline \text{Q3} \end{array} - \begin{array}{c} \text{108} \\ \hline \text{Q1} \end{array} = \begin{array}{c} \text{67} \\ \hline \text{四分位範圍} \end{array} \times 1.5 = \text{100.5}$$

最小値 (MIN)	8
第 1 四分位置 (Q 1)	108
中央値 (Q 2)	153
第 3 四分位置 (Q3)	175
最大値 (MAX)	290

$$\frac{175}{Q3} - \frac{108}{Q1} = \frac{67}{\text{四分位範囲}} \times 1.5 = 100.5$$

$$\frac{108}{Q1} = 100.5 = \frac{7.5}{\text{下のひげ}}$$

最小値 (MIN)	8
第 1 四分位置 (Q 1)	108
中央値 (Q 2)	153
第 3 四分位置 (Q3)	175
最大値 (MAX)	290

$$\frac{175}{Q3} - \frac{108}{Q1} = \frac{67}{\text{四分位範囲}} \times 1.5 = 100.5$$

$$\frac{108}{Q1} - 100.5 = 7.5$$

下のひげ

$$\frac{175}{Q3} + 100.5 = 275.5$$

上のひげ

外れ値ありの箱ヒゲ図は？

Group Activity

最小値 (MIN)	=QUARTILE(C3:C39,)	
第 1 四分位置 (Q 1)	QUARTILE(配列, 戻り値)	106
中央値 (Q 2)		152
第 3 四分位置 (Q3)		173
最大値 (MAX)		224

オプション

0 - 最小値

1 - 第 1 四分位数 (25%)

2 - 中位数 (50%)

3 - 第 3 四分位数 (75%)

4 - 最大値

1. Choose a world cup team:

<https://soccer.yahoo.co.jp/wcup/category/2022/teams>

The screenshot shows the 'FIFA WORLD CUP QATAR 2022' page on the Yahoo Japan sports website. The header features the tournament logo and navigation tabs: 'トップ' (Top), '日程・結果' (Schedule/Results), 'グループステージ' (Group Stage), 'ノックアウトステージ' (Knockout Stage), '個人成績' (Individual Results), 'チーム' (Teams), and '大会概要' (Tournament Overview). The 'チーム' tab is selected. On the left, a sidebar lists various sports categories, with 'サッカー' (Soccer) highlighted. The main content area displays a grid of 20 national teams, each with its flag, name, and links for '日程・結果' (Schedule/Results), '監督' (Coach), and '選手一覧' (Player List). The teams listed are: Japan, USA, Argentina, Iran, England, Wales, Uruguay, Ecuador, Australia, Netherlands, Ghana, Qatar, Canada, Cameroon, South Korea, Croatia, Costa Rica, and Saudi Arabia.

チーム	日程・結果	監督	選手一覧
日本	日程・結果	監督	選手一覧
アメリカ	日程・結果	監督	選手一覧
アルゼンチン	日程・結果	監督	選手一覧
イラン	日程・結果	監督	選手一覧
イングランド	日程・結果	監督	選手一覧
ウェールズ	日程・結果	監督	選手一覧
ウルグアイ	日程・結果	監督	選手一覧
エクアドル	日程・結果	監督	選手一覧
オーストラリア	日程・結果	監督	選手一覧
オランダ	日程・結果	監督	選手一覧
ガーナ	日程・結果	監督	選手一覧
カタール	日程・結果	監督	選手一覧
カナダ	日程・結果	監督	選手一覧
カメルーン	日程・結果	監督	選手一覧
韓国	日程・結果	監督	選手一覧
クロアチア	日程・結果	監督	選手一覧
コスタリカ	日程・結果	監督	選手一覧
サウジアラビア	日程・結果	監督	選手一覧

選手データを選択 ➡ Excelにコピー

特集	
W杯出場国戦力ランキング	
W杯選手ランキング	
サッカー小説 I'm BLUE	
球団別「期待の若手選手」	
種目	
プロ野球	
MLB	
高校野球	
侍ジャパン	
カタールW杯	
Jリーグ	
海外サッカー	
サッカー代表	
競馬	
ボートレース	
大相撲	
フィギュア	
格闘技	
ゴルフ	

日本 Japan

日程・結果監督・選手一覧

Pos.	背番号	選手名	所属チーム	生年月日	身長/体重	出身地
監督	-	森保 一	-	1968/8/23	-/-	長崎
GK	1	川島 永嗣	ストラスブール	1983/3/20	185/82	埼玉
GK	12	権田 修一	清水エスパルス	1989/3/3	187/84	東京
GK	23	シュニット ダニエル	シントトロイデン	1992/2/3	197/88	アメリカ
DF	2	山根 視来	川崎フロンターレ	1993/12/22	178/72	神奈川
DF	3	谷口 彰悟	川崎フロンターレ	1991/7/15	183/75	熊本
DF	4	板倉 滉	ポルシアMG	1997/1/27	186/75	神奈川
DF	5	長友 佑都	FC東京	1986/9/12	170/68	愛媛
DF	16	富安 健洋	アーセナル	1998/11/5	187/84	福岡
DF	19	酒井 宏樹	浦和レッズ	1990/4/12	185/78	長野
DF	22	吉田 麻也	シャルケ04	1988/8/24	189/87	長崎
DF	26	伊藤 洋輝	シュツットガルト	1999/5/12	186/78	静岡
MF	6	遠藤 航	シュツットガルト	1993/2/9	178/76	神奈川
MF	7	柴崎 岳	レガネス	1992/5/28	175/64	青森
MF	8	堂安 律	フライブルク	1998/6/16	172/70	兵庫

Excelで身長/体重の列をスプリット

Excelの「データ」タブの「区切り位置」ボタンをクリックして、区切り位置指定ウィザードを開きます。

ウィザードの「フィールドの区切り文字を指定してください。」画面で、区切り文字として「タブ」を選択します。

プレビュー画面で、身長/体重のデータが正しく区切りされていることを確認します。

完了ボタンをクリックして、区切り処理を完了します。

データタブ

区切り位置

区切り位置指定ウィザード - 2 / 3

フィールドの区切り文字を指定してください。

区切り文字

- ☒ タブ
- ☐ セミコロン
- ☐ カンマ
- ☐ スペース
- ☒ その他: /

☐ 連続した区切り文字は1文字として扱う

文字列の引用符: "

選択したデータのプレビュー:

身長	体重
-	-
185	82
187	84
197	88

キャンセル < 戻る 次へ > 完了

身長/体重列を選択

その他: /

所属チーム	生年月日	身長/体重
-	1968/8/23	-/-
ストラスブール	1983/3/20	185/82
清水エスパルス	1989/3/3	187/84
シントトロイ	1992/2/3	197/88
川崎フロンターレ	1993/12/22	178/72
川崎フロンターレ	1991/7/15	183/75
ボルシアMG	1997/1/27	186/75
FC東京	1986/9/12	170/68
アーセナル	1998/11/5	187/84
浦和レッズ	1990/4/12	185/78
シャルケ04	1988/8/24	175/70
シュツットガ르트	1999/5/12	173/68
シュツットガ르트	1993/2/9	173/68
レガネス	1992/5/28	173/68
フライブルグ	1998/6/16	172/70
ブライton	1995/1/16	174/68
モナコ	2001/6/4	173/67
ソシエダ	1995/5/10	177/74
スタッドラン	1993/3/9	176/66
フランクフルト	1996/8/5	180/72
デュッセルドルフ	1998/9/10	180/75
名古屋グランパス	1997/2/25	166/68
ボーフム	1994/11/10	173/71

その国の選手の次の5数要約を計算する：

0. 最小値
1. 第1四分位点
2. 中央値
3. 第3四分位点
4. 最大値

Draw a box and whisker diagram!