データとは

データとは、実験や調査、記録などによって得られた文字や数値・記号などのことです。
近年では、映像(えいぞう)や画像、音声などもデータとして扱(あつか)われることがあります。



データとは(データの種類)

数字で表されているデータは、以下のように分けることができます。

■名義尺度

出席番号・映画館や劇場での座席番号など、名前の代わりに他と区別するために存在する数のこと。



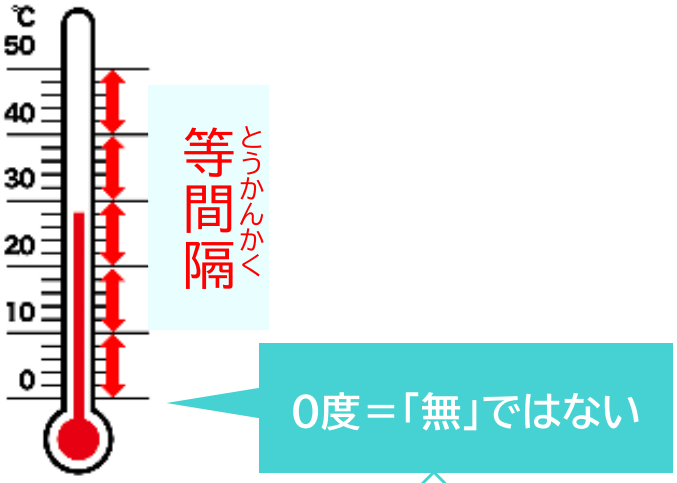
■順序尺度

順位(1位・2位・3位)や学年のように、数の大小に意味があるが、等間隔ではなく差や比率には意味がない数のこと。



■間隔尺度

気温や時刻のように、目盛りが等間隔で差に意味がある数のこと。ゼロにも意味がある。



■比例尺度:身長や金額のように、目盛りが等間隔で差や比率に意味がある数のこと。ゼロは「無い」ことを意味する。



名義尺度・順序尺度は「質的データ」

間隔尺度・比例尺度は「量的データ」と言われます。

データとは(データの種類)

クイズ！次のデータに用いられる尺度(しゃくど)を当ててみよう！！

①アイスクリームの種類



②来場者の数



③ぬいぐるみのサイズ



データとは(わたしたちの身近にあるデータ)

わたしたちの身近には、さまざまなデータがあります。

身近なところでは、天気(気温・降水量)、選挙の得票数や投票率などがあります。皆さんが住んでいる町の人口などは総務省統計局が行う国勢調査で調べられています。行政や企業活動など社会のあらゆる活動にデータが活かされており、わたしたちの生活の中でデータの重要性は年々大きくなっています。

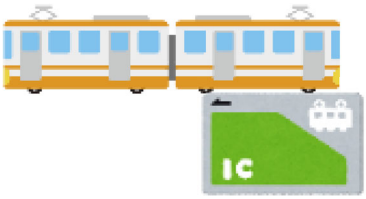




国勢調査とは

日本に居住しているすべての人を対象に、5年に一度実施する国の最も重要な統計調査で、性別・年齢・家族の人数・住居の種類・通っている学校や仕事のことなどあらゆることを調査しています。これらは、選挙を公平に行ったり、将来の人口の見通し、少子高齢化の予測など行政の運営に使われています。また、企業活動や大学の調査などにも役立てられています。

他にはこんなデータもあります



電車



カメラ
サーモグラフィー



YouTube
視聴履歴

データは生活のあらゆるところで収集されて

安全を守ったり、生活を便利にするために役立てられています

データとは(わたしたちの身近にあるデータ)

就業構造基本調査とは

日本国民の就業・不就業の状態を明らかにすることを目的として、統計法(国の統計に関する基本的な法律)に基づき実施する国の重要な統計調査です。5年ごとに実施される調査で、2022年10月に行われます。

就業構造基本調査では、正規・非正規雇用者の就業状況の違いや高齢層・若年層の就業状態、育児・介護と就業の関係などについてのさまざまな調査が行われ、雇用政策や経済政策などの企画・立案のための基礎資料として活用されます。

また、この調査結果はどなたでも閲覧・利用可能で、大学・研究機関等における学術研究などにも幅広く使われています。



**2022年10月1日に
調査が行われるよ**







就業構造基本調査の結果は、さまざまな取組に活かされています。

- 柔軟な働き方がしやすい環境整備
- 長時間労働の改善
- 職業能力の開発、人材育成
- 育児、介護・看護と就業の両立支援

など

データとは（社会における統計の活用）

わたしたちの身近にあるデータは、社会の中のさまざまな場面で役立っています。



【交通量のデータを使って渋滞を和らげる】

過去のデータから渋滞しそうな場所と時間帯を予測して交通渋滞が起こらないように役立てられています。



【雨雲レーダーを使って台風やゲリラ豪雨を予測】

雨雲のデータから台風やゲリラ豪雨を予測して、冠水など災害による被害が大きくならないように役立てられています。





【医療データを使って診断】

医療データは、診察時に過去の診断結果から病状を予測して、誤診を防ぐだけではなく病気の早期発見や予防につながるように役立てられています。



総務省統計局が国勢調査等の統計調査を実施しデータを提供することで、国や企業では、そのデータを皆さんの暮らしを良くするための活動に役立てています。具体的には、医療費や年金の給付額を決める・災害の対策に役立てる・企業が出店する場所を決めるときなどに使われます。

データの特徴を見てみよう

データの特徴をとらえやすくするためには、いくつかのポイントがあります。

ポイントが分かっていれば、たくさんのデータがあってもデータの特徴をとらえることができます。

次の表をぱっと見て、データの特徴は分かりますか？

図書館で借りた本の冊数

A組	5	3	8	2	15	4	9
B組	6	2	4	5	3	8	5



A組のほうが多そう
だけど‥

では、これではどうでしょうか？

図書館で借りた本の冊数

A組	2	3	1	2	15	4	9
B組	6	2	4	5	3	8	5
C組	4	7	15	2	1	3	16
D組	6	9	10	7	8	6	5
E組	11	2	3	20	4	1	5
F組	10	8	9	3	4	3	12
G組	8	15	4	7	2	6	3
H組	6	1	15	3	1	14	17
I組	16	12	10	7	17	2	1



見ただけでは
分からないよ･･･！！

たくさんのデータがあっても、データの特徴をとらえて見やすくするためのポイントをいくつか紹介します。

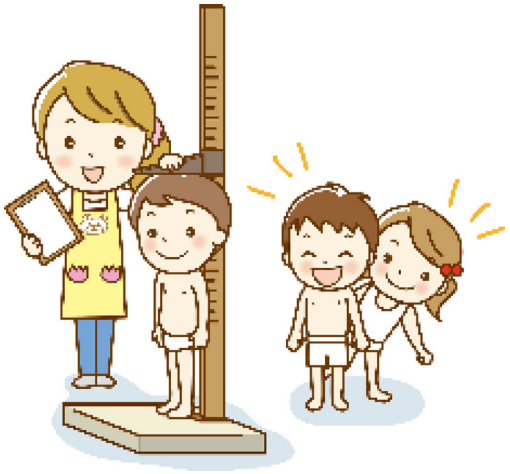
データの特徴(とくちょう)を見てみよう(平均)

平均とは、いくつかの数値に対して、それぞれをすべて足し合わせた後、数値の個数で割った値のことをいいます。

例えば7人の身長が以下のような場合、
7人の身長の合計値を人数で割った値が平均身長となります。

【例1】

名前	身長(cm)
Aさん	140
Bさん	135
Cさん	145
Dさん	155
Eさん	160
Fさん	150
Gさん	165



平均	(140+135+145+155+160+150+165)÷7=150(cm)

6人のテストの点数が以下のような場合、
6人の点数の合計値を人数で割った値が平均点となります。

【例2】

名前	点数(点)
Hさん	75
Iさん	80
Jさん	68
Kさん	96
Lさん	100
Mさん	61



平均	(75+80+68+96+100+61)÷6=80(点)

データの特徴を見てみよう（中央値）

平均の他にも、データの特徴を見るために重要な値はいくつかあります。中央値は、データを小さい順に並べたときに真ん中に位置する数字のことを言います。データが5個ある場合は、小さい順に並べたときに真ん中に来る、3番目の数字が中央値になります。データが6個ある場合は、小さい順に並べたときに真ん中に来る、3番目と4番目の平均が中央値です。





中央値はいくつ？

中央値を求めてみよう



データの特徴を見てみよう（最頻値）

最頻値は、データの中でもっとも登場回数が多い値のことを言います。

以下の例の場合、数値（値段）ごとの登場回数を求めて、もっとも登場回数が多い値が最頻値です。

もっとも登場回数が多い値が2個以上ある場合は、それら全てが最頻値です。

（例1）

メニュー	値段（円）
かっぱ巻き	100
納豆巻き	100
たまご	100
いなり寿司	100
いか	150
サーモン	150
たこ	150
えんがわ	200
えび	200
しめさば	300
あなご	400
たい	400

値段（円）	度数
100	4
150	3
200	2
300	1
400	2

最頻値は 100

（例2）

メニュー	値段（円）
かっぱ巻き	100
納豆巻き	100
たまご	100
いなり寿司	100
いか	150
サーモン	150
たこ	150
赤貝	150
えんがわ	200
えび	200
しめさば	300
あなご	400
たい	400

値段（円）	度数
100	4
150	4
200	2
300	1
400	2

最頻値は
100と150



「代表値」

ここまでで説明した「平均」「中央値」「最頻値」などの値を「代表値」と呼びます。
代表値には以下のような種類があります。

平均　中央値　最頻値　最大値　最小値