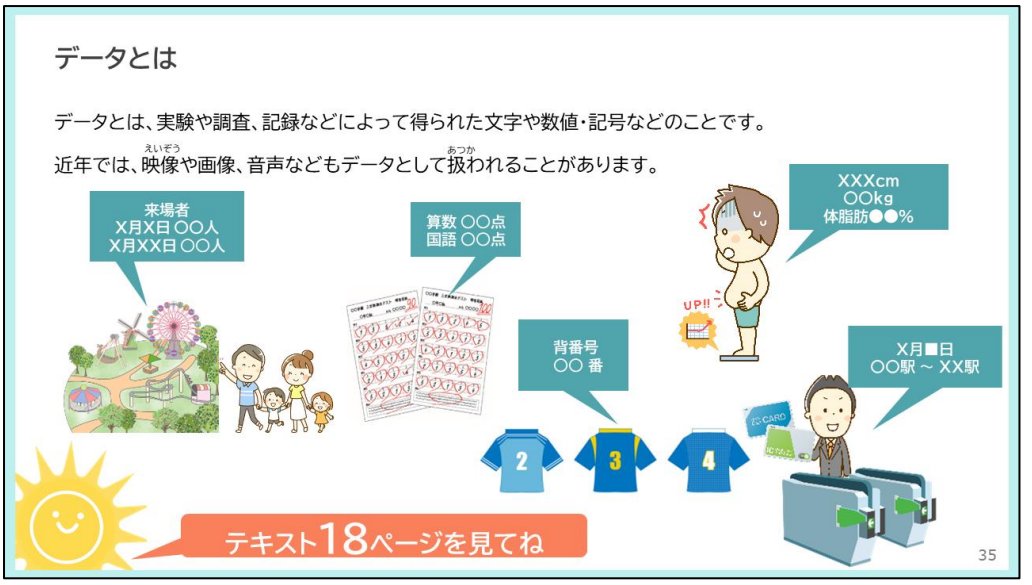
休　憩　中

休憩が終わったら
データについて説明するよ！

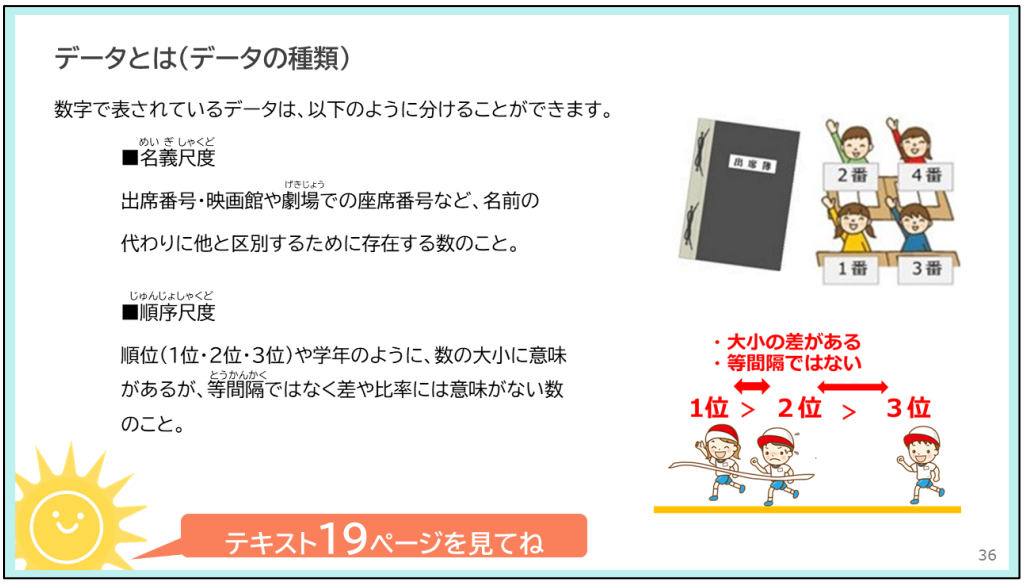




【スライドの簡単な解説】
さまざまな形式で記録された情報を、総合して「データ」といいます。成績表や店の売上金額などのように数字で表わすことができるもの以外にも、動画配信サイトで視聴ができるような「音声」や「映像」などもデータとして扱われます。

【このページのポイント】
・紙に書かれた情報は、コンピュータが読み取れる状態に電子化することで「データ」と呼ばれるようになります
・数字、記号のほか文字や画像、映像なども「データ」として扱われます
・日常生活のいたるところで、データは取得されています（駅の改札、身体測定など）

【キーワード】
音声データ、画像データ、映像データ



【スライドの簡単な解説】
データは、その役割によって4つに分類されます。このページでは、そのうち2つを説明します。
①名義尺度（めいぎしゃくど）：分類するためだけにつけられたデータ（出席番号、座席番号など）
②順序尺度（じゅんじょしゃくど）：順番に意味をもつが、データ同士の距離・間隔には意味をもたないデータ（順位・学年など）

【このページのポイント】
順序尺度は、次のページに出てくる間隔尺度と間違えやすいので要注意！
数字の順番に意味があり、数字同士の差に意味がないものを「順序尺度」、数字の順番にも差にも意味があるものを「間隔尺度」と覚えましょう。

【キーワード】
名義尺度、順序尺度



【スライドの簡単な解説】
1つ前のページで紹介した4つの分類のうち、残りの2つを説明します。

③間隔尺度（かんかくしゃくど）：温度が1℃から2℃に上昇したときに、1℃から2倍になったとは言いません。温度の1℃と2℃は差のみに意味があります。温度にはマイナスがあるので、0℃にも意味があります。
④比例尺度（ひれいしゃくど）：身長、年齢のように、差と比率の両方に意味があります。たとえば、身長が100㎝から150㎝になったとき、身長が1.5倍になったと言うことができます。さらに0（ゼロ）が「無い」を意味します。（身長0cm、金額0円=無い）

【このページのポイント】
「間隔尺度」「比例尺度」とも目盛りが等間隔という点は同じですが、0（ゼロ）に意味があるかどうかで区別することが出来ます。

【キーワード】
間隔尺度、比例尺度



【スライドの簡単な解説】
前のページで学習した尺度についてのクイズです。

【このページのポイント】
①はサイズの大小に意味はあるかないかに注目して考えてみましょう。
②と③はゼロに意味があるかないかに注目して考えてみましょう。



【スライドの簡単な解説】
尺度当てクイズの答えを解説します。

【このページのポイント】
①のアイスクリームの種類は、分類するためだけにつけられた名前です。よって正解は名義尺度です。

②の来場者の数は、もし来場者数か0人だった場合。来場者は「いない」ということになります。よって正解は比例尺度です。

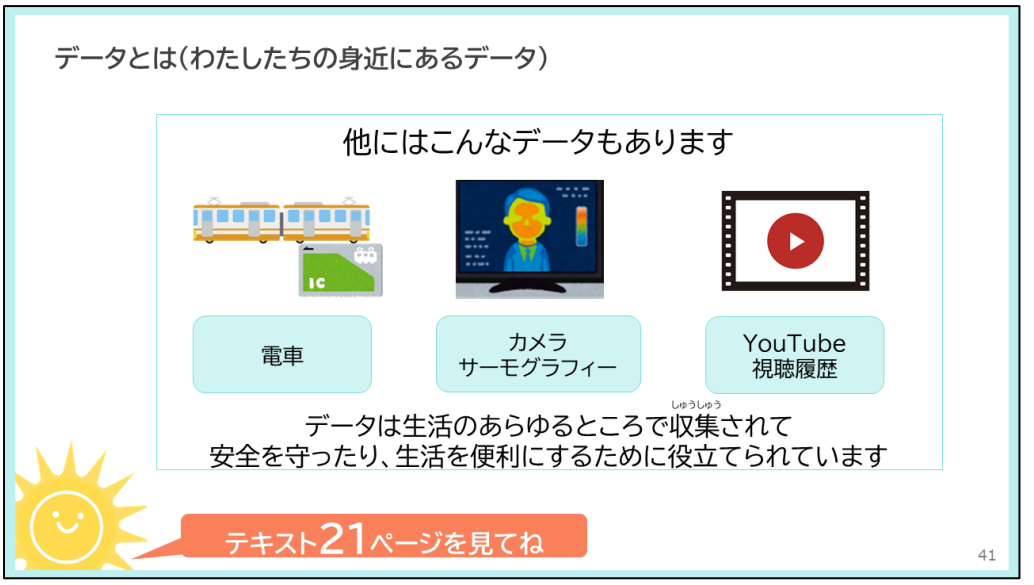
③のぬいぐるみのサイズは、S・M・Lと大きさの大小に意味あることがわかります。よって正解は順序尺度です。



【スライドの簡単な解説】
国勢調査は、5年に1回行われる日本全国に住んでいる世帯(家族)全員を対象にした調査です。みなさんの生活について調査を行うことで、より住みやすい国づくりに役立てられます。

【このページのポイント】
国勢調査で集められたデータは、私たちの日常生活における住みよいまちづくり、防災対策などに活用されています。

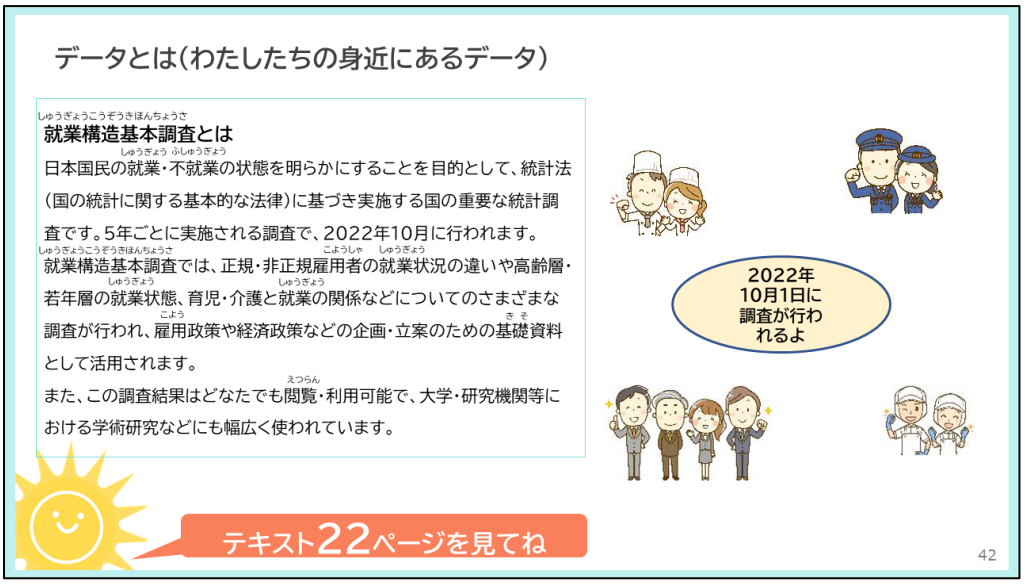
【キーワード】
国勢調査（こくせいちょうさ）



【スライドの簡単な解説】
現代は、ありとあらゆるものが「データ」として取得されています。交通機関や街中の映像、動画配信サービスなどもそのひとつです。
動画配信サービスなどでは、「ログ」と呼ばれるいつ誰が何を見たかのデータを用いて計算し、個々人がそれぞれ好む動画を推測し、次に見る動画としておすすめするシステムなどが開発され、運用されています。

【このページのポイント】
私たちの身近には、人のあらゆる行動の情報をデータとして取得する仕組みがたくさんあります。

【キーワード】
行動データ＝ログデータ



【スライドの簡単な解説】

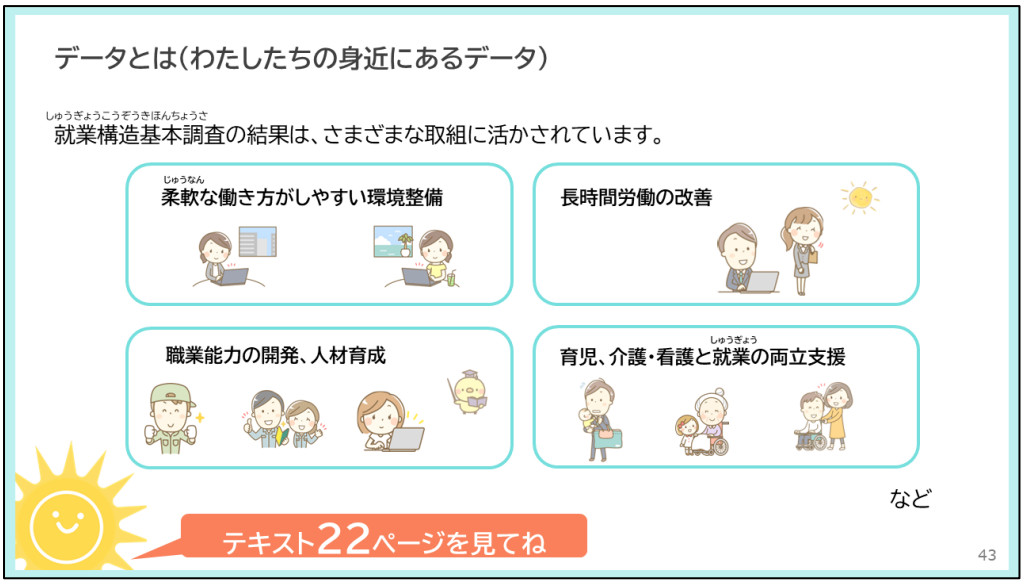
就業構造基本調査とは、5年ごとに行われる、日本国内に住む人が仕事についているかどうかの調査です。日本国内に住む人の仕事について、全国及び地域別、年齢別などといったより詳細な状況を明らかにすることを目的として行われています。

【このページのポイント】

総務省統計局はさまざまな調査を行い、得られたデータを整理して公表しています。その一つが就業構造基本調査です。

【キーワード】

総務省統計局、就業構造基本調査

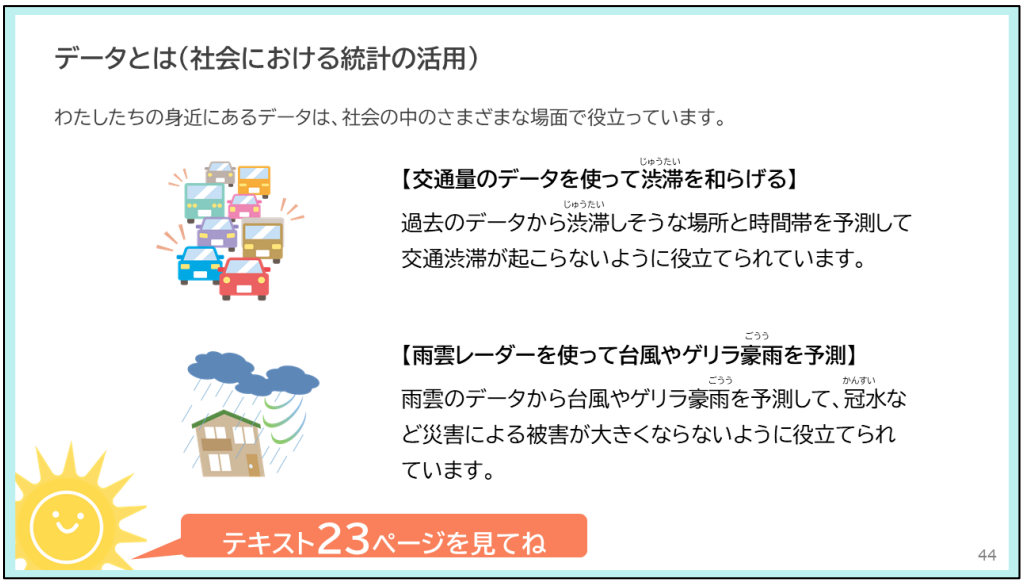


【スライドの簡単な解説】

就業構造基本調査の結果は、国の基本的な方針決定の基礎資料として活用されるほか、地方公共団体における雇用対策などの各種施策に利用されています。たとえば、就業構造基本調査の結果から仕事をする女性が増えていることが分かれば、仕事をする女性が仕事と育児や介護などを両立できるように，柔軟な働き方ができる環境を整える施策を提案できます。

【このページのポイント】

総務省統計局が行う就業構造基本調査の結果は、国や地方自治体などの施策に活かされています。

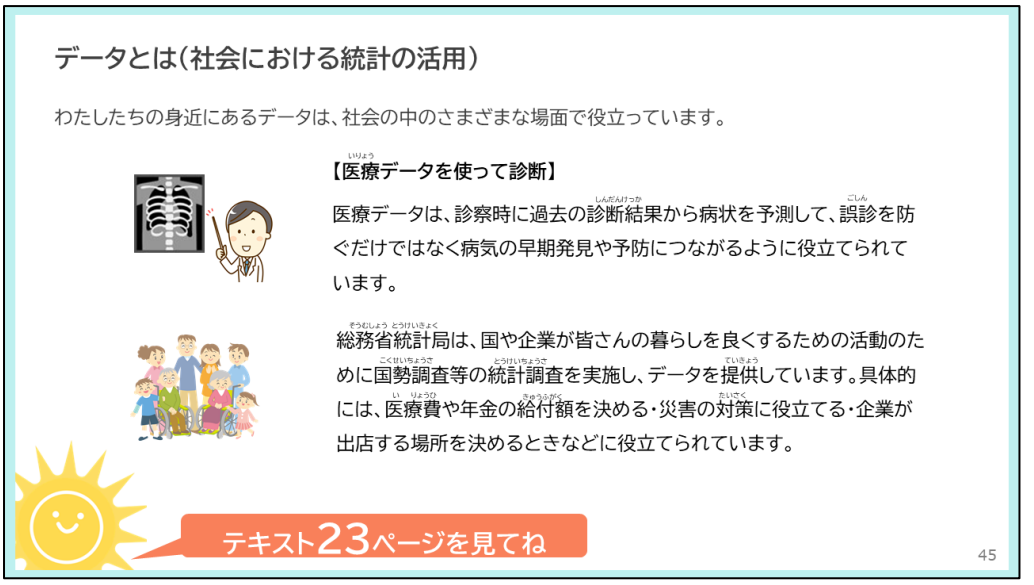


【スライドの簡単な解説】

生活の中で使われるデータで、特にみなさんに身近なものとして「交通」「天気」に関するデータがあります。

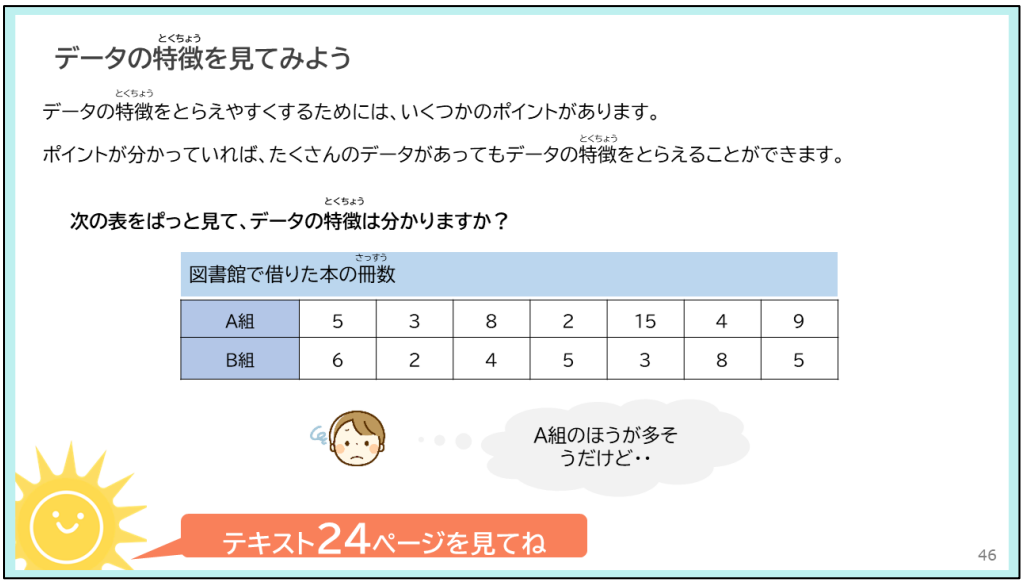
交通渋滞を防ぐために、過去に取得したデータが活用されています。

雨雲データからゲリラ豪雨を予測し、早めの避難の呼びかけに活用されています。



【スライドの簡単な解説】
病院などにおいても、病気の診断をするために過去のデータが活用されています。
たとえば、過去に癌（がん）になった人の情報は、癌（がん）の早期発見や予防、治療に役立てられています。

【このページのポイント】
データを集めることで、未来の人々の健康や安全を守るために活用することができるのです。



【スライドの簡単な解説】
データの取り扱い方についての解説。A組とB組のこどもたちが図書館で借りた本の冊数を比較してみて、どんな特徴があるのかを考えます。

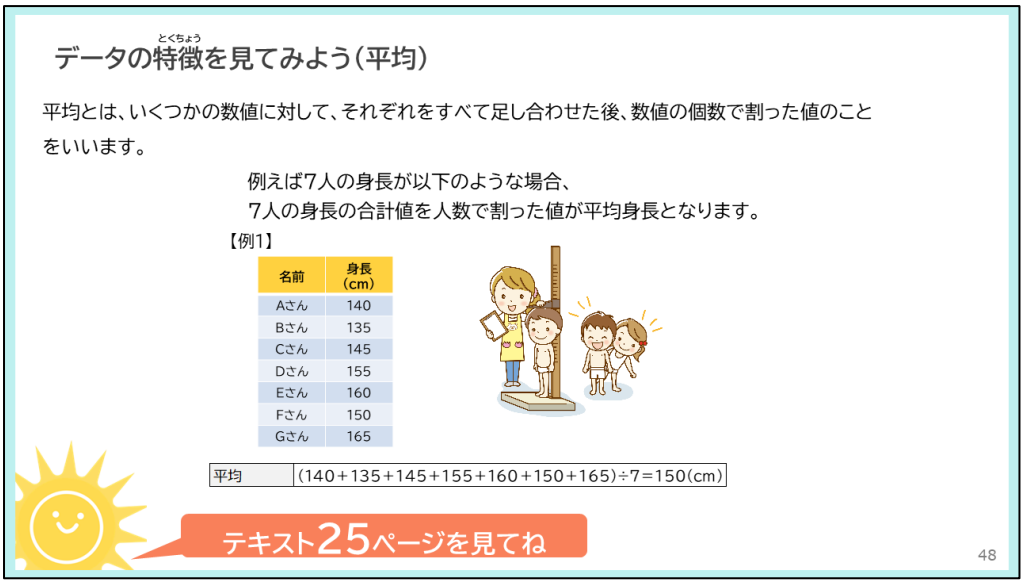
【このページのポイント】
A組とB組のこどもたちが図書館で借りた本の冊数を比較した表を見ると、ぱっと見てA組の借りた本の数が多いことがわかります。
データの量が少ないと、表の数字だけを見ただけでも特徴が把握できます。



【スライドの簡単な解説】
データの取り扱い方についての解説。A組からI組の9クラスのこどもたちが図書館で借りた本の冊数を比較してみて、どんな特徴があるのかを考えます。

【このページのポイント】
A組からI組の9クラスのこどもたちが図書館で借りた本の冊数を比較したときに、すぐに特徴をとらえることは難しいことが分かります。
交通渋滞の予測や、医療の画像から癌（がん）かどうかを診断するような時にはたくさんのデータが必要となります。
このようなビッグデータから特徴をすぐに把握するためのポイントがいくつかあります。

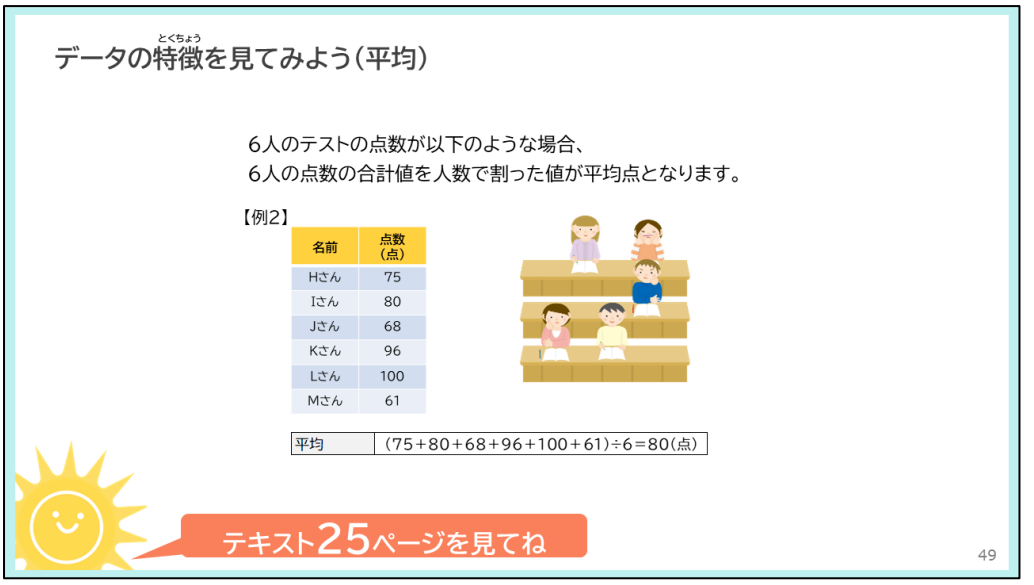
【キーワード】
ビッグデータ…世の中にあるたくさんのデータ



【スライドの簡単な解説】
たくさんのデータから特徴をとらえる方法の１つである「平均」についての解説。

【このページのポイント】
このページでは身長の平均を求めます。表にのっているAさんからGさんの身長を足し合わせ、それを人数(7人)で割った値が平均身長となります。

【キーワード】
平均…いくつかの数値を全部足して、その個数で割ったもの



【スライドの簡単な解説】
たくさんのデータから特徴をとらえる方法の１つである「平均」についての解説。

【このページのポイント】
このページでは表のテストの点数の平均を求めます。表に示されているHさんからMさんの点数を全部足し合わせ、それを人数(６人)で割った値が平均点となります。

【キーワード】
平均…いくつかの数値を全部足して、その個数で割ったもの



【スライドの簡単な解説】
たくさんのデータから特徴をとらえる方法の１つである「中央値」についての解説。

【このページのポイント】
図の上段には５個のお寿司が値段の安いものから高いものへと順に並んでいます。この場合の中央値は真ん中に位置するエビの200円です。
また、データが偶数個の場合は真ん中に位置する２つを足して２で割った値が中央値となります。
下段のように６個のお寿司が安い順に並んでいる場合は、真ん中に位置する３番目(エビ200円)と４番目(タイ400円) を足した数値(値段)を２で割った数値が中央値です。
つまり、(200＋400)÷2＝300となるので中央値は300円です。

【キーワード】
中央値…データを小さい順から大きい順に、または大きい順から小さい順に並べたときの真ん中に位置する値をいう。



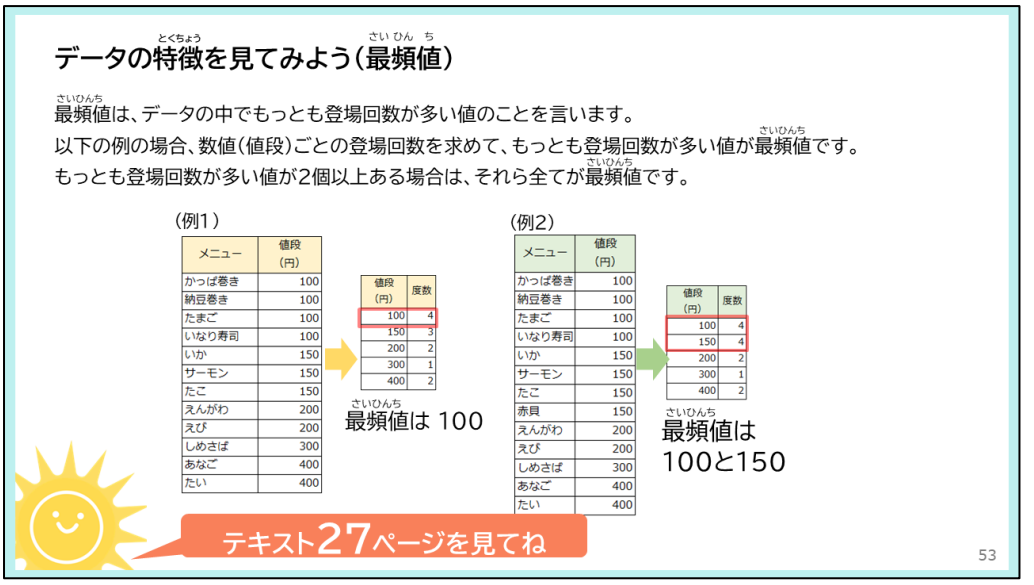
【スライドの簡単な解説】
(演習)「中央値」を求めよう。

【このページのポイント】
お寿司の数が6個（偶数個）であることに注目してみましょう。



【スライドの簡単な解説】
「中央値」の問題の答えを解説します。

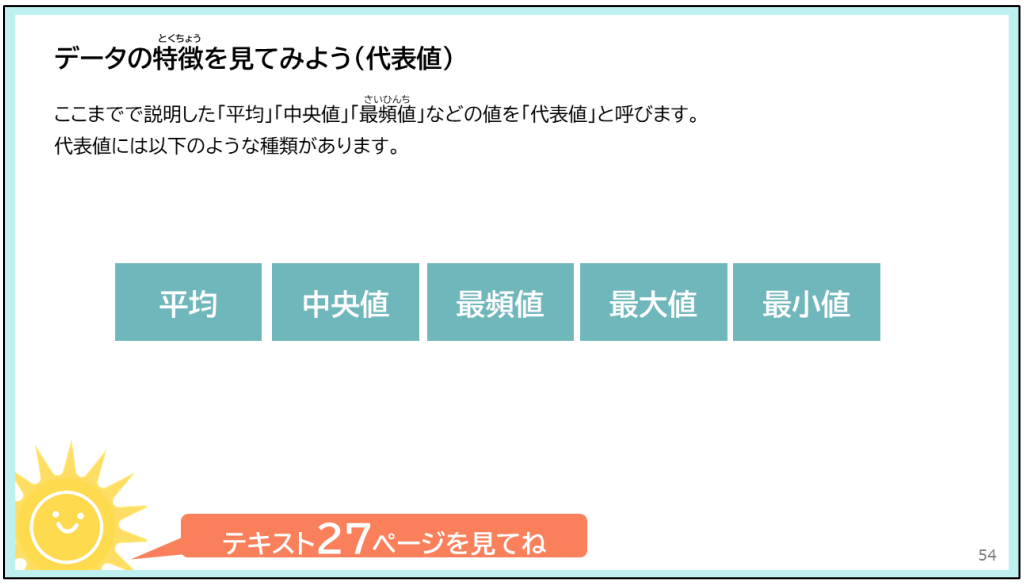
【このページのポイント】
まず、中央値を求めるには小さいものから大きいものへと順に並べます。お寿司の数が偶数ですから真ん中に位置するお寿司はタコとうなぎの２個です。タコ250円とうなぎ450円を足して2で割ると中央値がでます。
つまり、(250＋450)÷2＝350となるので、中央値は350円です。



【スライドの簡単な解説】
たくさんのデータから特徴をとらえる方法の1つである「最頻値」についての解説。

【このページのポイント】
メニューに登場する回数がもっとも多い値段のお寿司を調べます。
(例1)では、100円のお寿司の値段が登場回数がもっとも多いことがわかります。
つまり、ここでの最頻値は100円です。
(例2)では、100円と150円のお寿司の値段が登場回数がもっとも多いことがわかります。
登場回数がもっとも多いお寿司の値段が２つの場合には、どちらも最頻値といえるので、最頻値は100円と150円となります。

【キーワード】
最頻値…データの中でもっとも登場回数が多い値



【スライドの簡単な解説】代表値の解説

【このページのポイント】
「平均」、「中央値」、「最頻値」の3つはどのようなデータを見るときも必ずチェックします。
今後さらに統計学やビッグデータについて学習していくと、ここに書いてある「最小値」や「最大値」もチェックするようになります。
「平均」、「中央値」、「最頻値」の3つに「最小値」や「最大値」を加えたものを、データを代表する値として「代表値」といいます。

【キーワード】
最大値…データの中でもっとも大きい値。
最小値…データの中でもっとも小さい値。
代表値…データの特徴をひとつの数値として表す「平均」「中央値」「最頻値」などの総称。