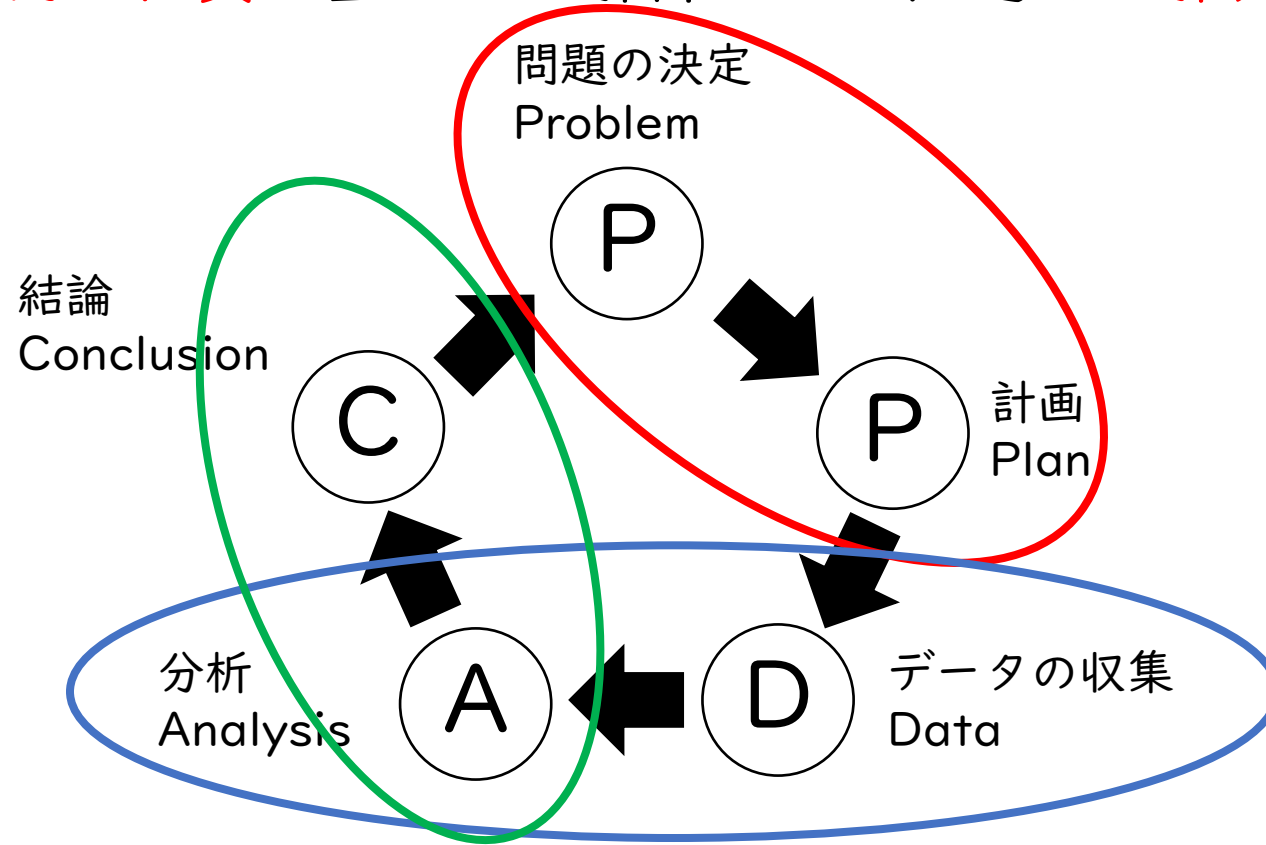


データサイエンス入門

なぜ今、データサイエンスに注目が集まっているのか。
学習する目的やその必要性、重要性について理解する。

データサイエンスとは

- データサイエンスとは 【引用：日本大百科全書（ニッポニカ）】
 - さまざまな課題の解決や展望を予測するため、
 - 膨大に蓄積されているデータの内容やその分布を調べ、
 - 特定の傾向や性質に基づいた解析により、適切な解決方法を提示・評価する



一度で課題をすべて
解決できるのは稀



結果から新たな課題
に取り組む

(再掲) データサイエンスとは

- データサイエンスとは 【引用：日本大百科全書（ニッポニカ）】
 - さまざまな課題の解決や展望を予測するため、
 - 膨大に蓄積されているデータの内容やその分布を調べ、
 - 特定の傾向や性質に基づいた解析により、適切な解決方法を提示・評価する
- データサイエンスは「データ」が無いと始まらない
 - データは世の中にたくさんあふれている
 - データが散在していると、手が付けられない
 - まずはデータをまとめるところからスタート

データをまとめる・蓄積する①

- データにも様々な種類があることに気づく
 - 数値（整数、実数、カテゴリーのラベルなど）
 - 文字（出現回数をカウントすると数値データ）
 - 画像や音

- 一つの物に対して、
様々なデータがある
例：お菓子の表示

商品名: [REDACTED] ●名称:清涼菓子 ●原材料名:甘味料(ソルビトール、アスパルテーム・L-フェニルアラニン化合物)、香料、微粒酸化ケイ素、ショ糖エステル、クチナシ色素、(一部にゼラチンを含む)
●内容量:50粒(7g) ●賞味期限:左下部に年月記載 ●保存方法:高温多湿を避けて保存
●販売者: [REDACTED]

栄養成分表示
1製品(7g)当たり

エネルギー	23kcal
たんぱく質	0.01~0.2g
脂質	0.04~0.4g
炭水化物	6.7g
糖類	0g
食塩相当量	0g

4 946842 501908

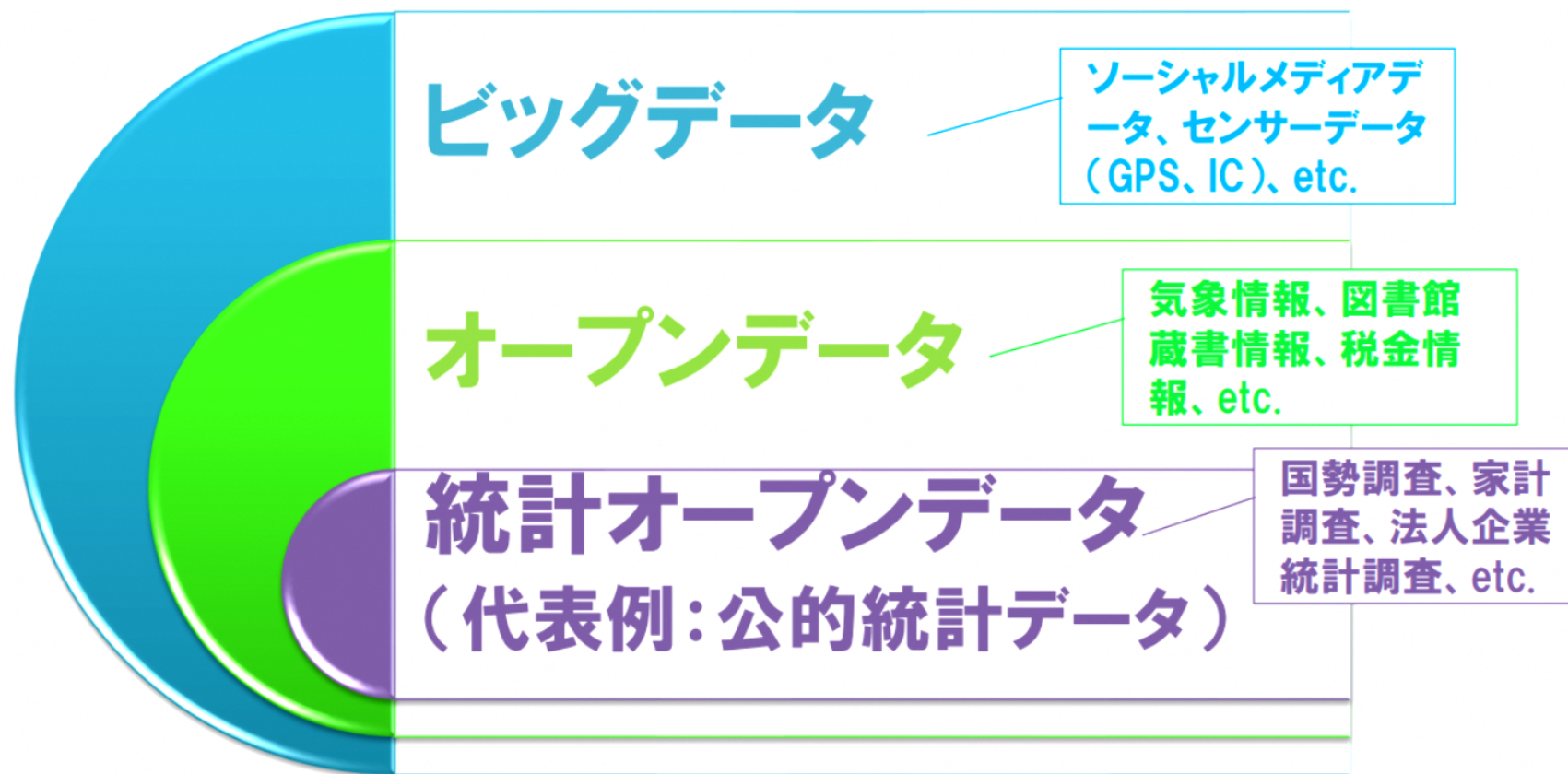
2025.11/06 7XE

218

データをまとめる・蓄積する②

- データにも様々な種類があることに気づく
 - 数値（整数、実数、カテゴリーのラベルなど）
 - 文字（出現回数をカウントすると数値データ）
 - 画像や音
- 一つの物に対して、様々なデータがある
 - （目的に合わせて）必要な項目を抽出し、それをまとめる
 - これらが蓄積してくるとデータベースが充実してくる
（ただし、実際にデータベースは縁の下の力持ち的な役割で外からは見えにくい）

データが集まったものの分類

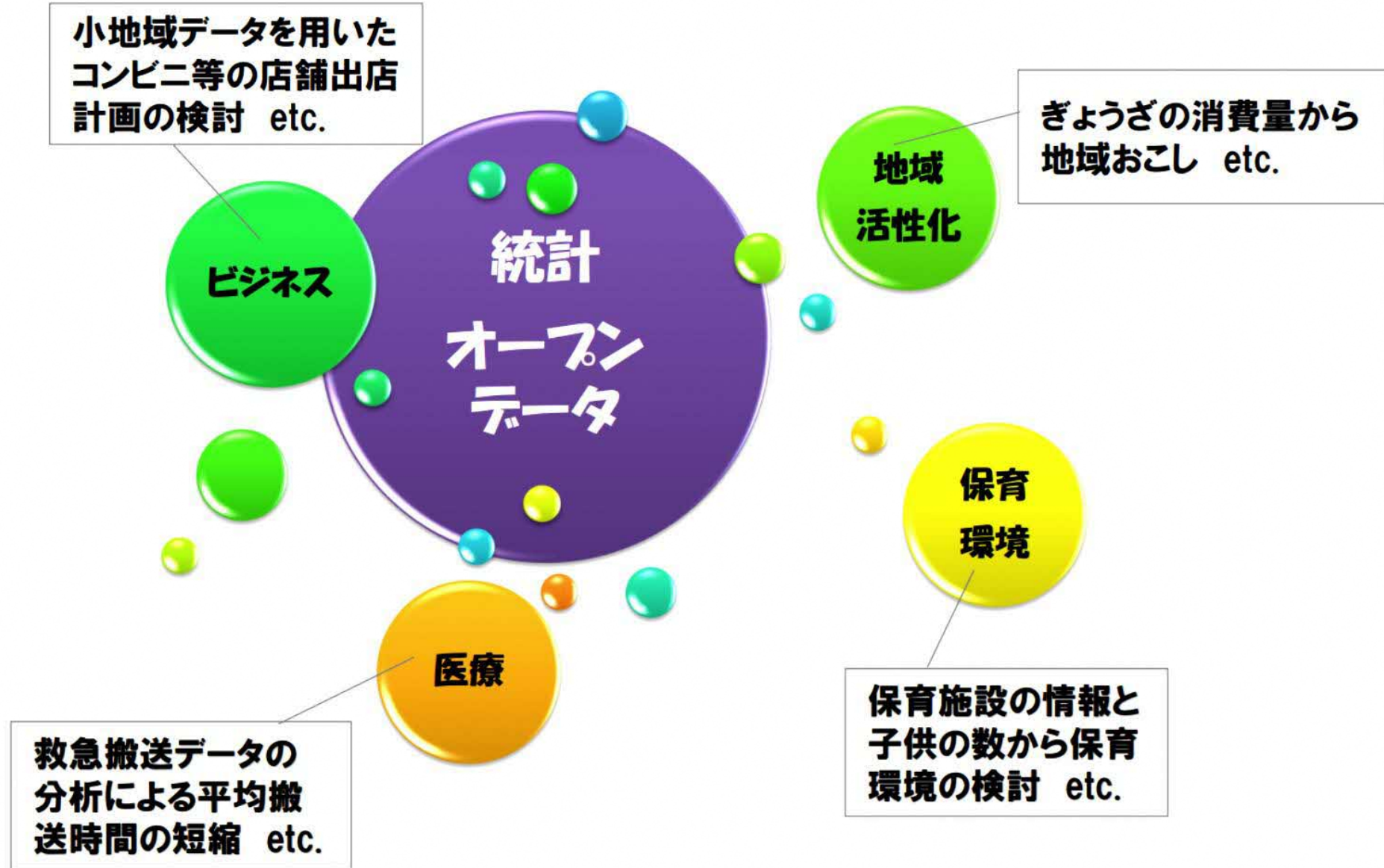


ビッグデータ... ICT (情報通信技術) の進展により生成・収集・蓄積等が可能・容易となる多種多様のデータ

オープンデータ... 機械判読 (利用) 可能な形態で、二次利用を認める条件の下、ネット上公開されるデータ (情報)

統計オープンデータ... 公的統計は、ビッグデータの中の代表的なオープンデータ

例えば、統計オープンデータの活用



データが集まれば、いよいよ分析

ただ、第一印象は

- 「データ分析は難しそう・・・」
- 「今までやったことないのに、できるのか不安・・・」
- 「データ分析って何をどうすればよいの？」

という、思いがあるのでは

実は第1歩目がけっこう大変？

- まず何から手を付けたらよいか？『分析の目的を明確にする』
 - 普段の生活で感じる「なぜ？」「どうして？」「本当だろうか？」に対して、理由や真偽をはっきりさせたいことを**目的**として設定（**仮説**を設定）
- 目的を達成するためにどうやって価値あるデータを集めたらよいか？
 - 実はここが一番大変だったり（難しかったり）する
 - 例：生徒からアンケートをとることが実行しやすいと思うが・・・
（質問内容や対象人数等、工夫や試行錯誤が必要）

実践事例：「アンケートで身近な問題を解決しよう！」

<https://www.youtube.com/watch?v=VFzCUPp30jg>

(参考) データを集め整理すること

「データを集計するだけで
様々な考察を行うことができる」

クイズ形式で
考えてみましょう



クイズ

コンビニでおでんが最も売れる時期は何月？

A.
9～10月

B.
11～12月

C.
1～2月

D.
3～4月

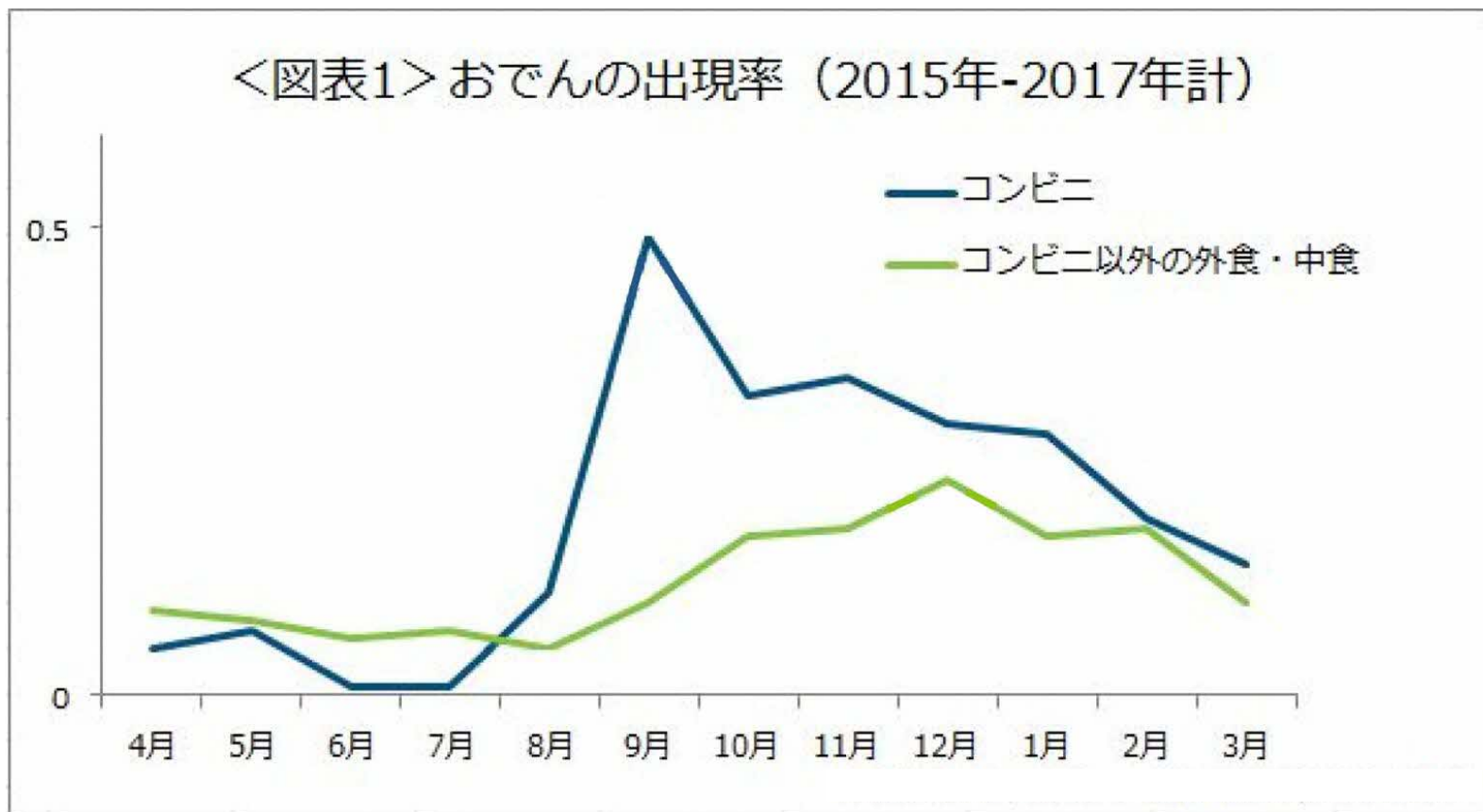
答え

クイズ

コンビニでおでんが最も売れる時期は何月？

A.
9～10月

コンビニでおでんが最も売れる時期は何月？



おでんの出現率（2015年 - 2017年計） Circana, サカーナ・ジャパン調べ

仮説 1

「おでんは寒い時」というイメージ



- ・人は寒暖差によって「寒さ」を感じる
- ・もっとも寒さを感じるのが10月

※朝晩の寒暖差5度以上

仮説 2

コンビニ各社は9月上旬頃には
おでんを宣伝するテレビCMを放映



購買意欲をそそる

続いて、
次のクイズ



Lampone

Yogurt

Pesca

Bacio

クイズ

日本で一番アイスクリームを食べている都市は？
(一世帯あたり年間支出金額1位)

A.
那覇市

B.
札幌市

C.
金沢市

D.
福井市

答え

クイズ

アイスクリームを食べている都市は？
(一世帯あたり年間支出金額1位)

C.
金沢市

アイスクリームの年間購入額

都市名	年間購入金額 (円)	順位 (位)
金沢市	12,828	1
さいたま市	12,710	2
福島市	11,632	3
山形市	11,222	4
千葉市	11,107	5
高知市	11,035	6
富山市	11,012	7
熊本市	11,006	8
宇都宮市	10,962	9
盛岡市	10,883	10

出典：家計調査（二人以上の世帯）品目別都道府県庁所在市及び政令指定都市ランキング 2020年（令和2年）～2022年（令和4年）平均（総務省統計局）
（<https://www.stat.go.jp/data/kakei/rank/backnumber.html>）を基にライフイズテック株式会社作成



さらに考えてみよう

なんで？

仮説



加賀百万石の繁栄時代国が豊か
だったため文化が栄えた



和菓子文化をはじめとする
甘味嗜好が色濃いのでは？

仮説検証

ケーキの
年間購入額Top5

都市名	年間購入金額 (円)	順位 (位)
金沢市	9,746	1
さいたま市	9,352	2
富山市	8,958	3
東京都区部	8,893	4
千葉市	8,872	5

チョコレートの
年間購入額Top5

都市名	年間購入金額 (円)	順位 (位)
金沢市	7,986	1
岐阜市	7,839	2
大津市	7,757	3
鳥取市	7,692	4
札幌市	7,678	5

まんじゅう・ようかんを除く
和生菓子の年間購入額Top5

都市名	年間購入金額 (円)	順位 (位)
金沢市	12,600	1
岐阜市	12,557	2
山形市	11,778	3
奈良市	11,270	4
津市	10,824	5

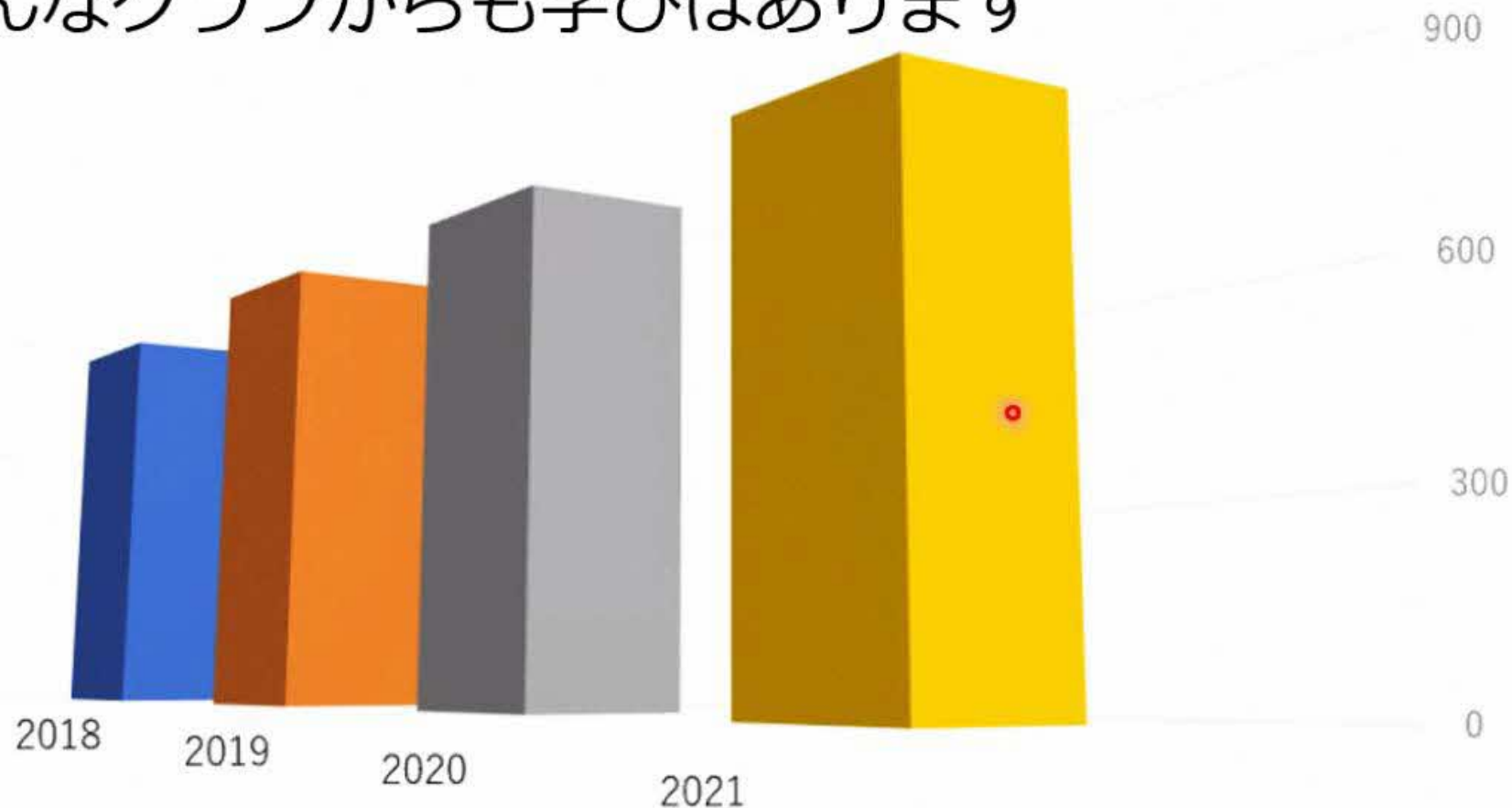
<再掲>

「データを集計するだけで
様々な考察を行うことができる」

→ 「集め整理することは重要」

(参考) 簡単な例からみると①

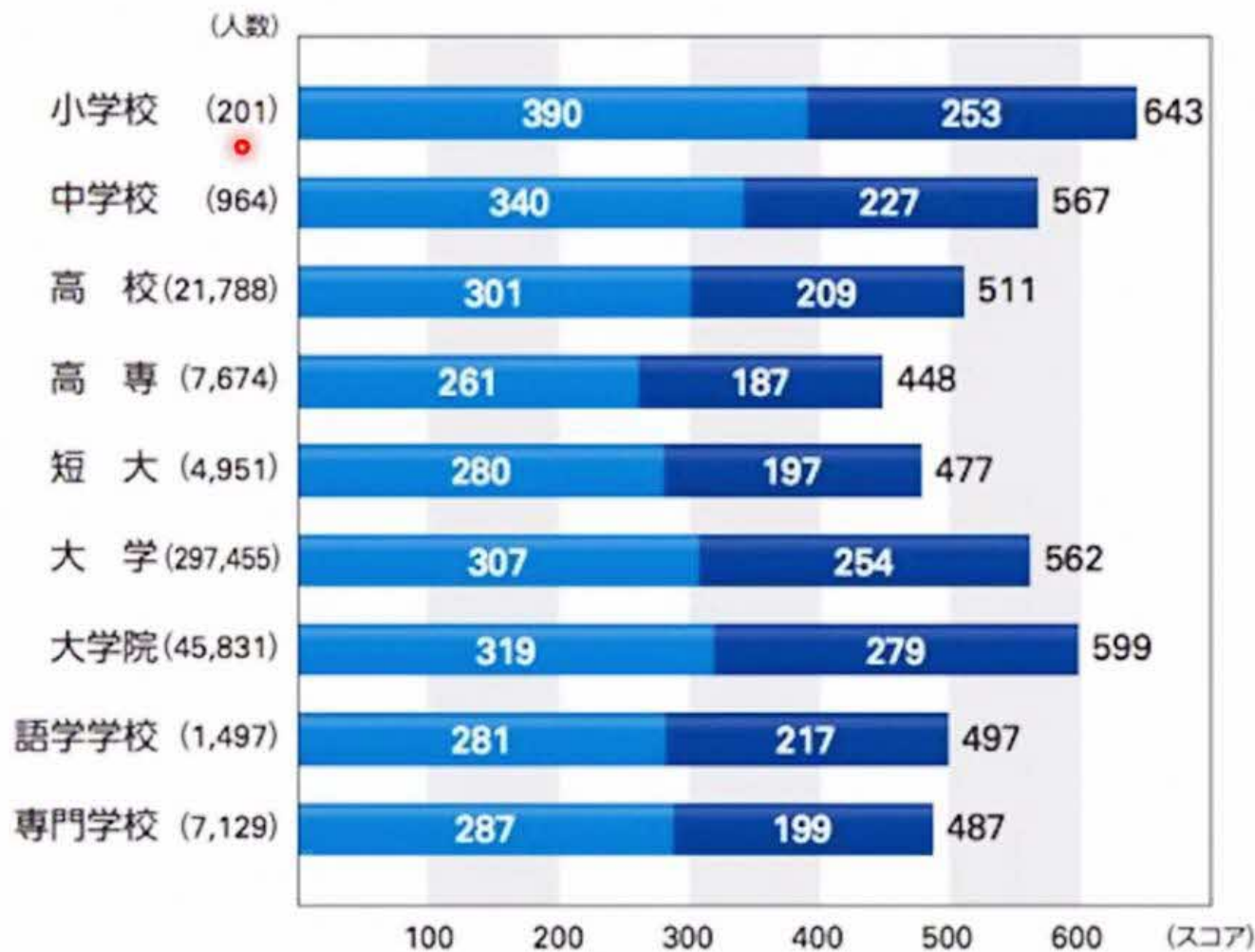
例：こんなグラフからも学びはあります



2018	2019	2020	2021
823	846	840	830

(参考) 簡単な例からみると②

例：TOEICは小学生が一番よくできる？



母集団は？
無作為抽出？

こういったデータの
各項目に関する
特徴を見ることが
意外と重要

(参考) <https://www.procrastist.com/entry/science-literacy>

(参考) 新しい職業も
例えば『データサイエンティスト』



データサイエンティスト
ってなに？

最後に

○統計学の基本を学習します。

○アンケート調査の基本を学習します。

○実務に活かす統計学を学習します。

知識やスキルの獲得のみにとどまらず、

コンピテンシーにつながるように意識して取り組んでください。