

ゼロからはじめる統計学

第1回

「統計学の歴史／要約統計量」



講師紹介

- 氏名

- 岡崎 凌（おかざき りょう）



- 学歴

- 大阪大学 修士課程修了（工学）

- 産業科学研究所

- 次世代半導体の研究開発

- 半導体特性の画像解析のために機械学習を導入

- 職歴

- Teaching Assistant(大学教員補佐)

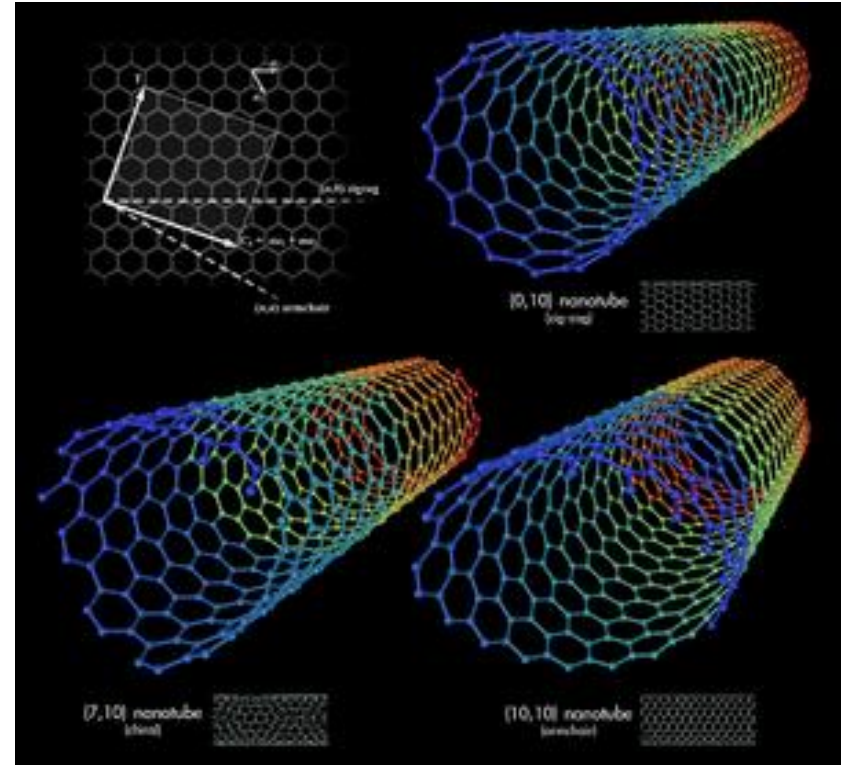
- 企業研修

- 趣味

- サッカー

- ピアノ

- 数学（数学オリンピック銀賞）



- スキル

- ・プログラミングによる環境構築(C言語, Java)
 - ・機械学習（Deep Learningによる特徴量抽出）
 - ・ニューラルネットワーク

統計学の必要性

- データはたくさんあるのに…
- 感覚で行なっている判断を数字を使って行いたい
- 個人のスキルに頼らない判断が行えるようになりたい



ビッグデータ・AI

データの保管、蓄積が容易になり、高度な分析も手軽に

新技術、仕事にも影響

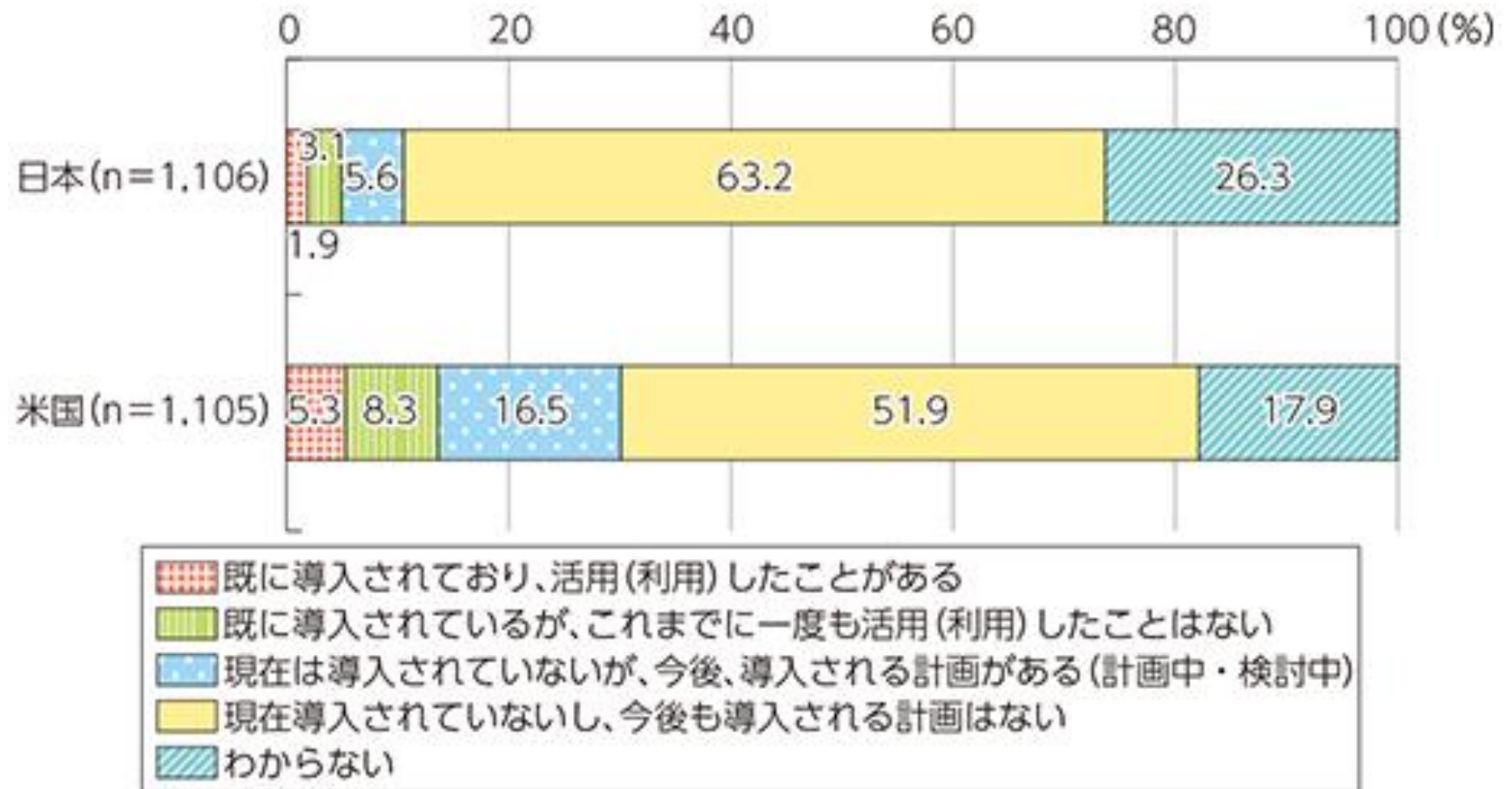


T-POINT



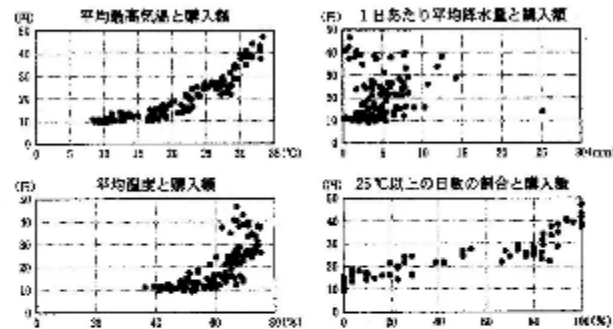
ビッグデータ・AI

- 会社でも取り入れようとする傾向



「データ分析」が必修化

〔2〕 次の4つの散布図は、2003年から2012年までの120か月の東京の月別データをまとめたものである。それぞれ、1日の最高気温の月平均(以下、平均最高気温)、1日あたり平均降水量、平均湿度、最高気温25℃以上の日数の割合を横軸にとり、各世帯の1日あたりアイスクリーム平均購入額(以下、購入額)を縦軸としてある。



出典：総務省統計局(2013)「家計調査年報」、『過去の気象データ』(気象庁Webページ)などにより作成

次の ☐ ス、☐ セ に当てはまるものを、下の①～④のうちから一つずつ選べ。ただし、解答の順序は問わない。

これらの散布図から読み取れることとして正しいものは、☒ スと ☐ セ である。

- ① 平均最高気温が高くなるほど購入額は増加する傾向がある。
- ② 1日あたり平均降水量が多くなるほど購入額は増加する傾向がある。
- ③ 平均湿度が高くなるほど購入額の散らばりは小さくなる傾向がある。
- ④ 25℃以上の日数の割合が80%未満の月は、購入額が30円を超えていない。
- ⑤ この中で正の相関があるのは、平均湿度と購入額の図のみである。

(数学I・数学A第2問は次ページに続く。)

2015年より「データ分析」が必修化

- 「データ分析」では、
- データの整理と分析
 - 色々な種類のグラフの読み取りが問われる

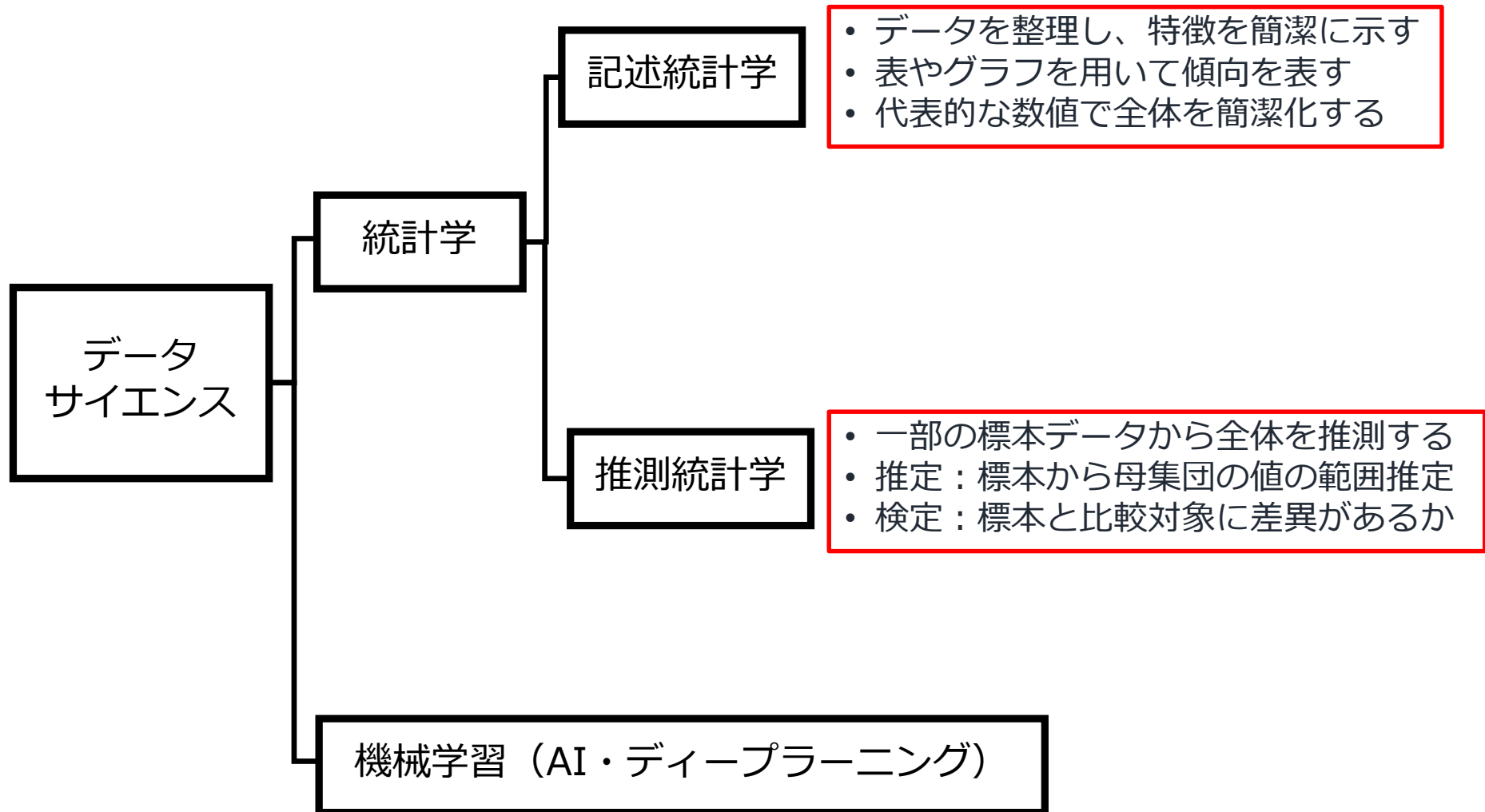
(センター試験「数学I・A」より)

AI人材年25万人育成へ



(日経新聞:政府、AI人材年25万人育成へ 全大学生に初級教育)

データ分析マップ



記述統計

データの集計

ID	満足度	他者評価	プロジェクト数	労働時間 (月平均)	労働時間 (会社内)	Work accident	退職・在職	過去5年の 昇進	所属部署	給料
1019	0.36	0.47	2	136	3	0	退職	無	accountin g	low
6830	0.68	0.51	5	158	3	0	在職	無	technical	medium
9653	0.53	0.64	2	109	3	0	在職	無	hr	medium
12208	0.78	0.87	4	228	5	0	退職	無	support	low
4816	0.92	0.56	4	170	3	0	在職	無	marketin g	medium

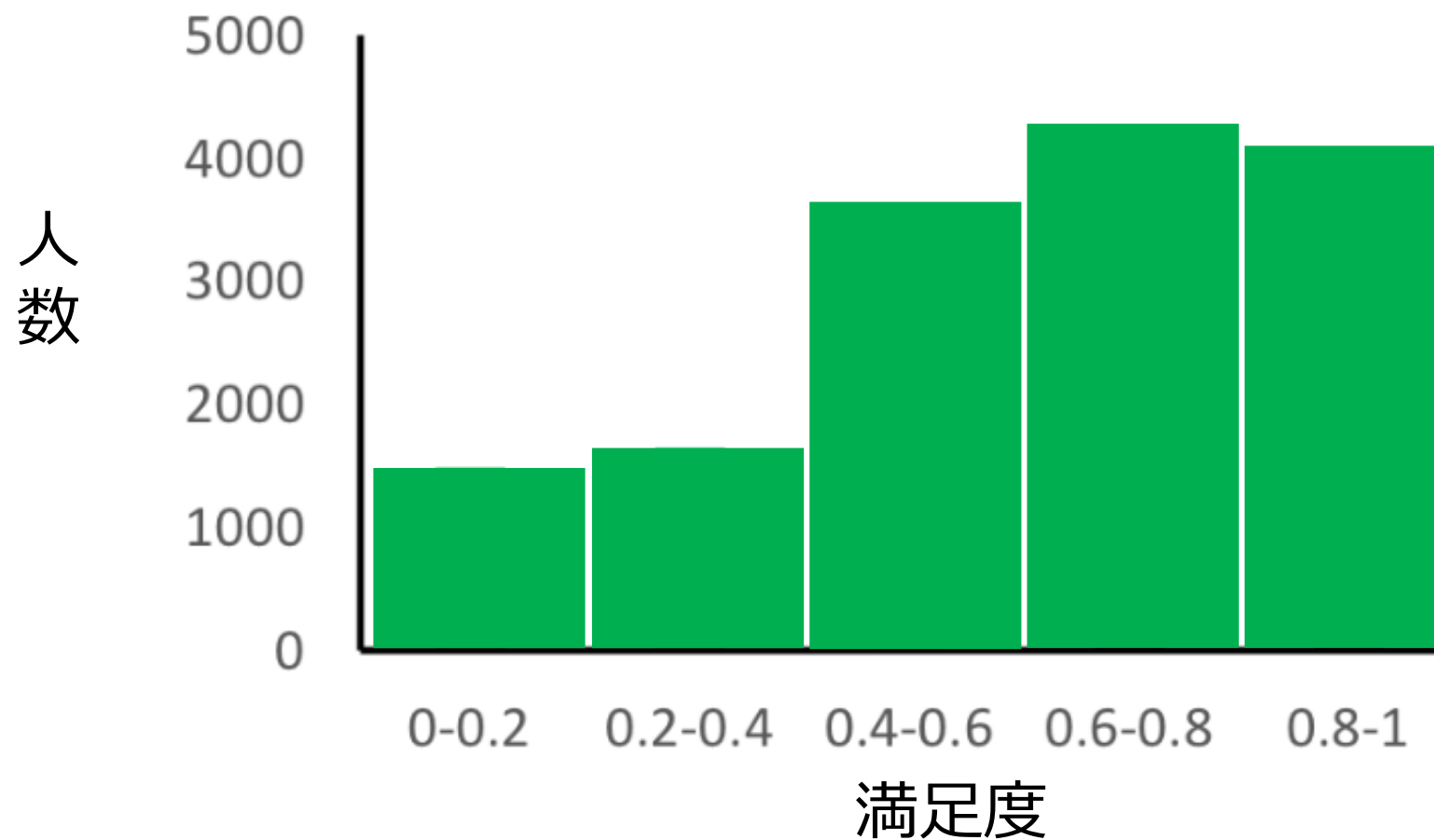
記述統計

データの要約

ID	満足度	他者評価	プロジェクト数	労働時間 (月平均)	労働時間 (会社内)	Work accident	退職・在職	過去5年の 昇進	所属部署	給料
1019	0.36	0.47	2	136	3	0	退職	無	accountin g	low
6830	0.68	0.51	5	158	3	0	在職	無	technical	medium
9653	0.53	0.64	2	109	3	0	在職	無	hr	medium
12208	0.78	0.87	4	228	5	0	退職	無	support	low
4816	0.92	0.56	4	170	3	0	在職	無	marketin g	medium

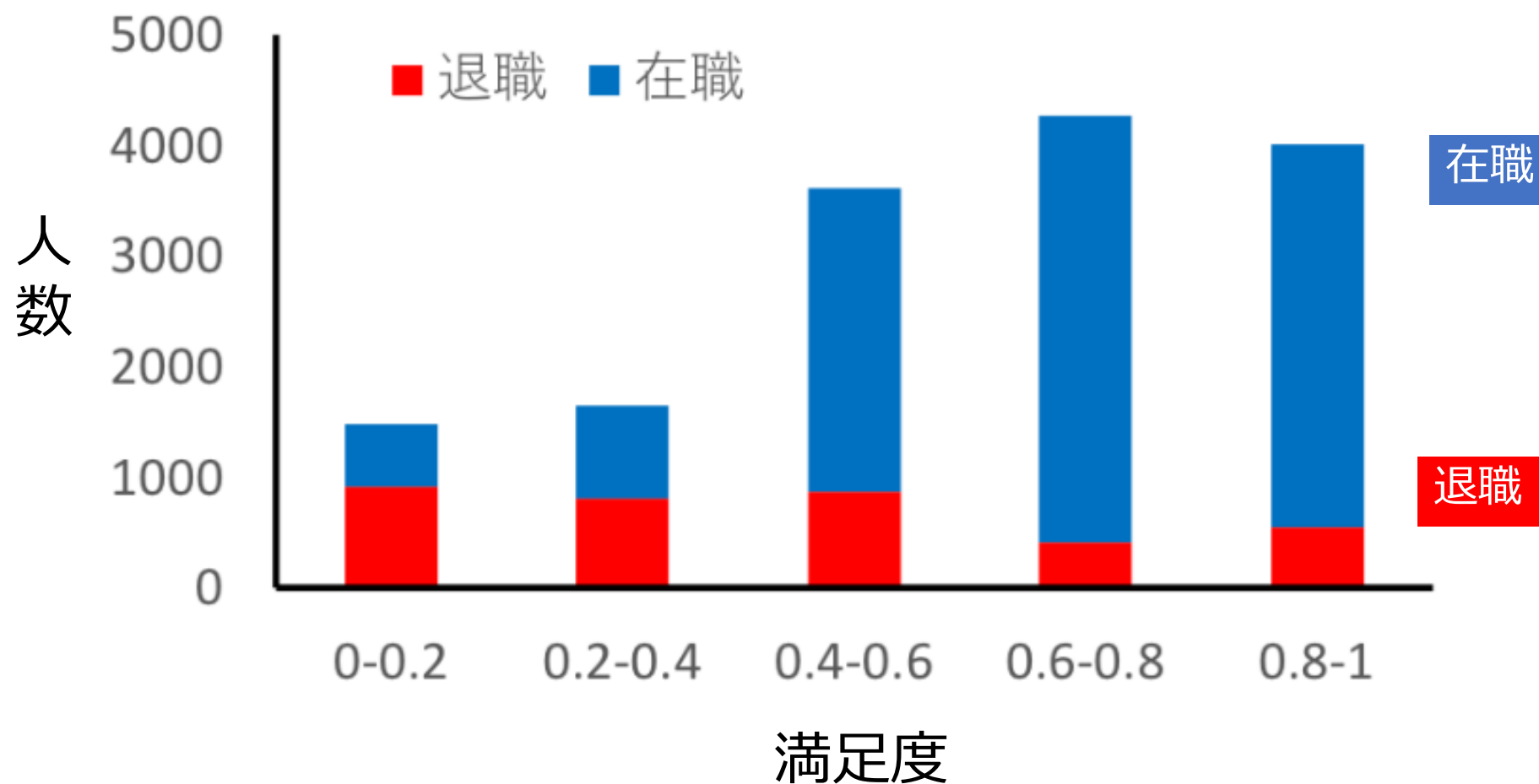
記述統計

データの可視化



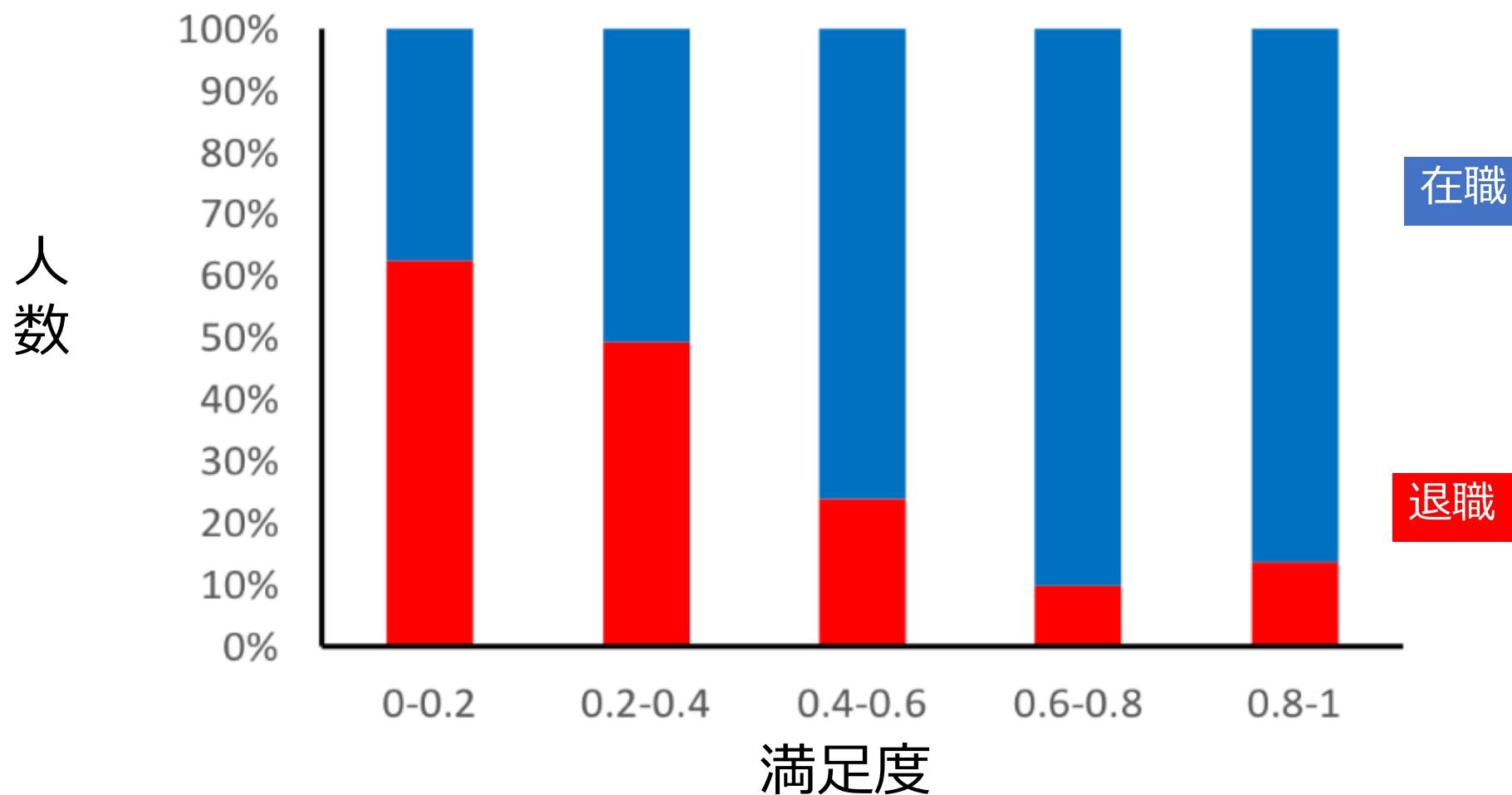
記述統計

データの可視化



記述統計

データの可視化



推測統計

- ・ 確率論をもとに、予測の信頼性を検証する

カイ二乗検定（新薬の効果検証）



	新薬	標準薬
有効性あり	25	44
有効性なし	5	12
計	30	56

推測統計学

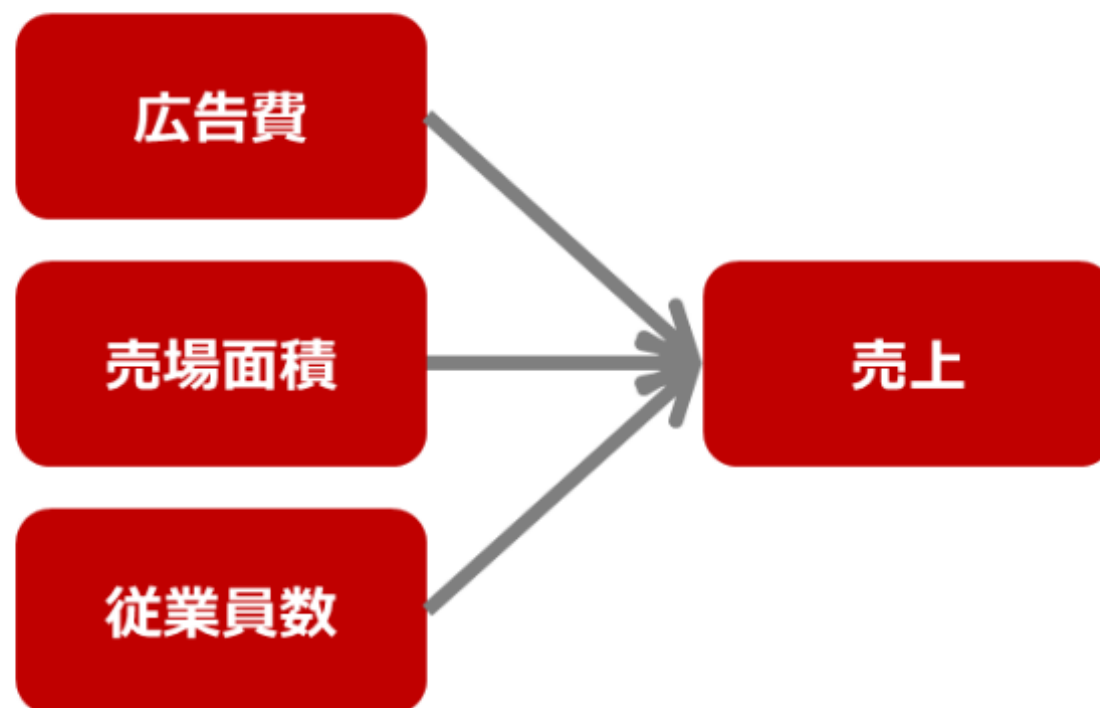
- 確率論を利用して、データの関係性を調べて予測を行う
t検定（ABテスト：Webデザインの比較）



推測統計学

- 確率論を利用して、データの関係性を調べて予測を行う

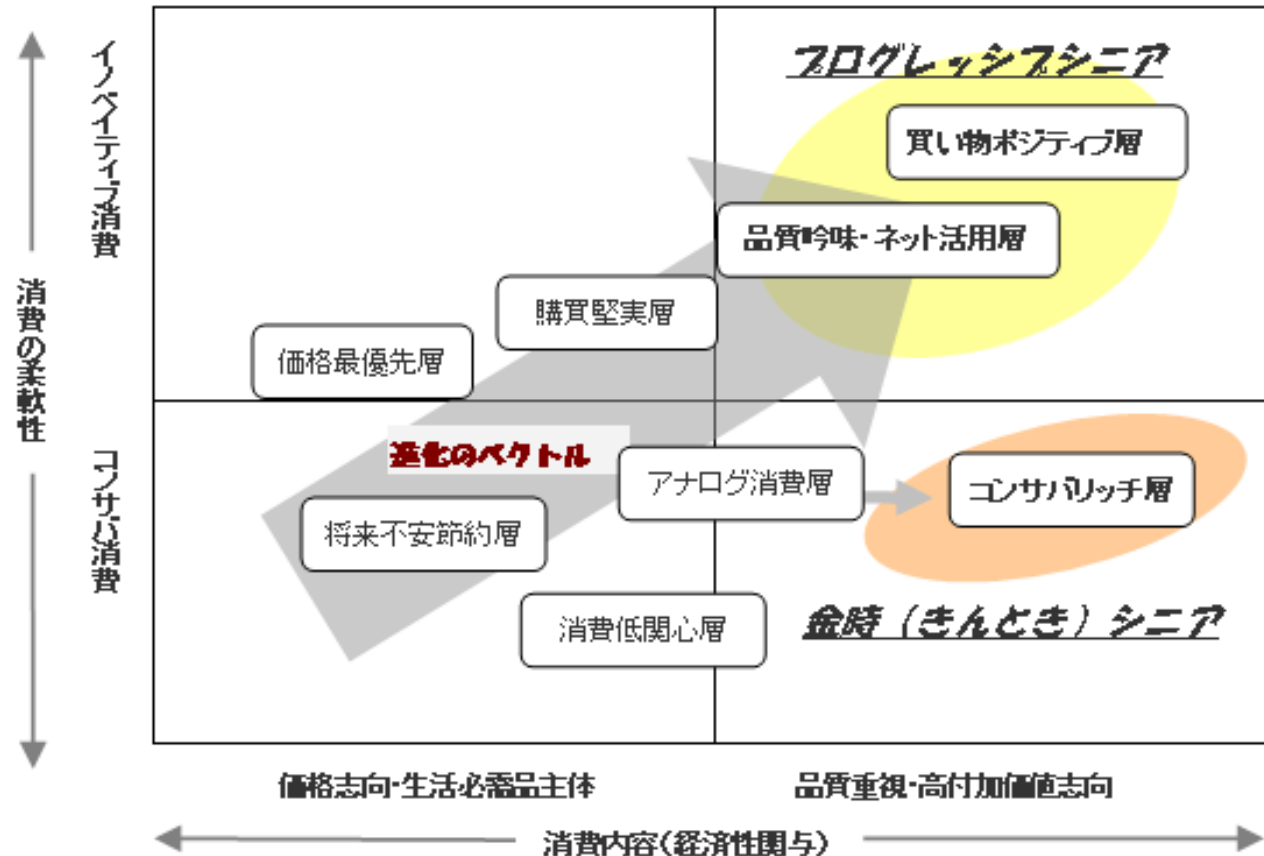
回帰分析（売上の予測モデル）



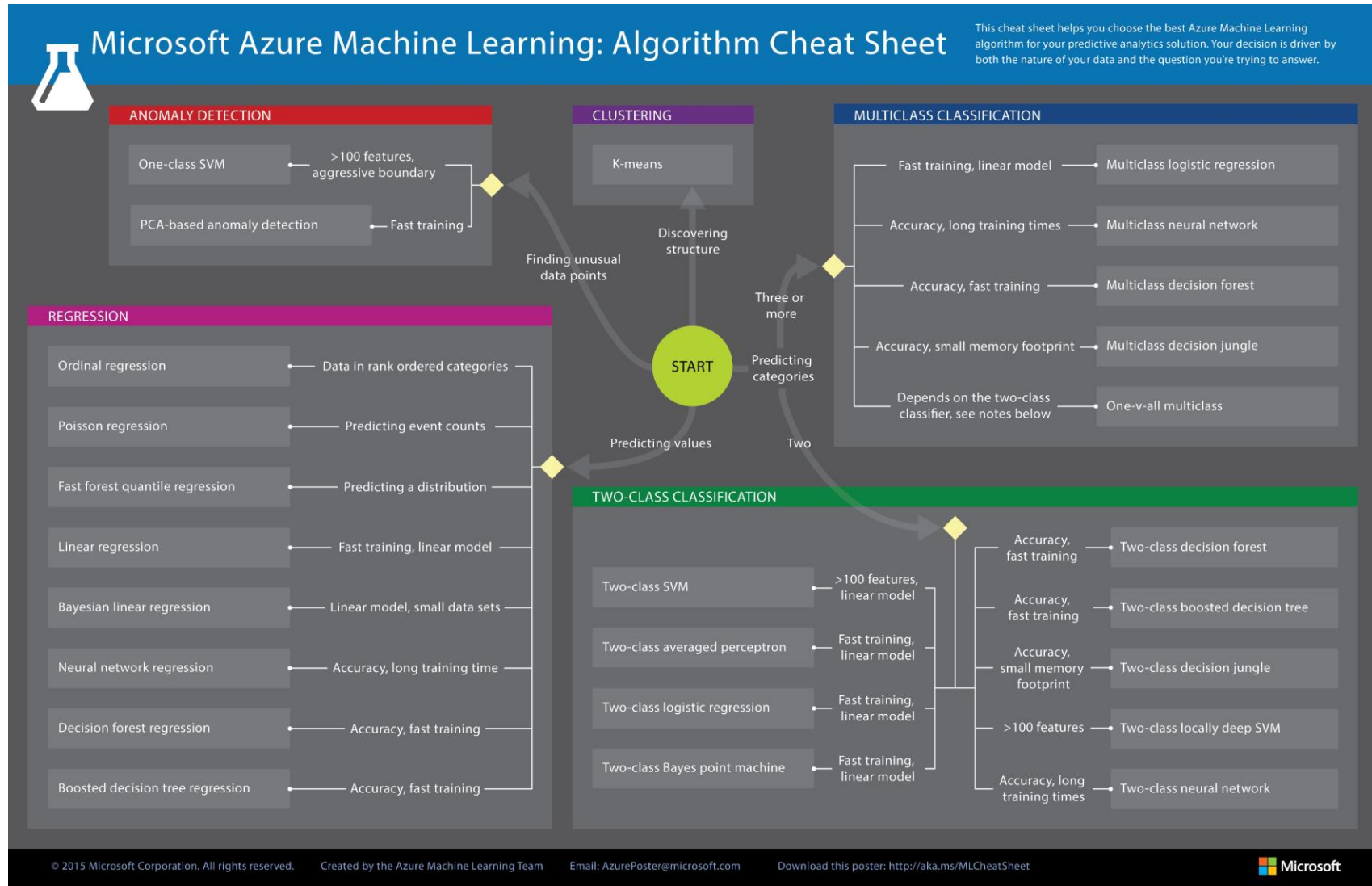
多変量解析

- 大量のデータの中で、傾向を考える
クラスタ分析（顧客のグループ化）

【図 2: シニアの消費行動クラスターの階層構造】(コレスポネンス分析)



機械学習



Deep Learning(画像認識)

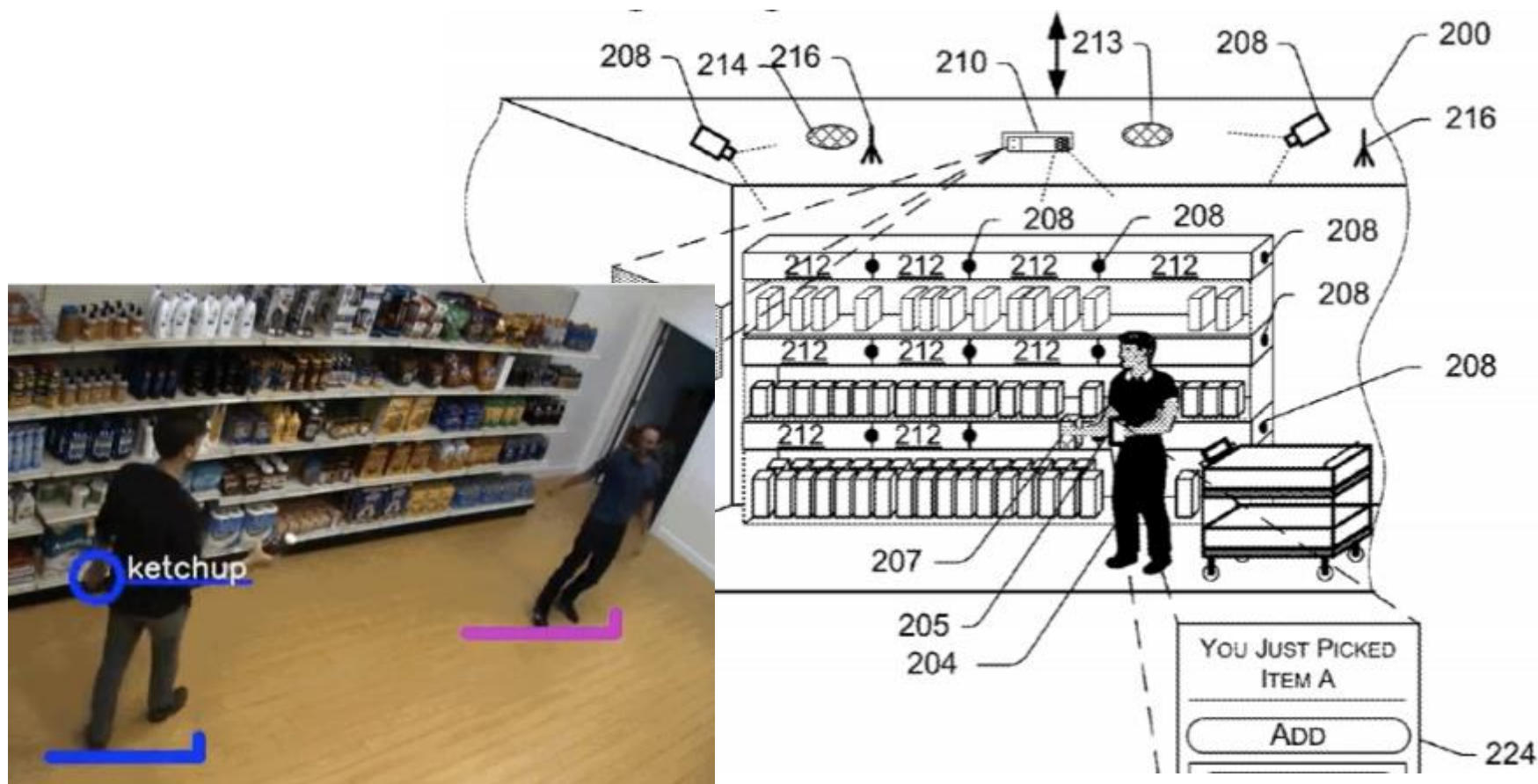
2018年、アメリカで無人店舗「Amazon Go」 開店



<https://youtu.be/NrmMk1Myrxc>

Deep Learning(画像認識)

2018年、アメリカで無人店舗「Amazon Go」 開店



統計学の歴史

統計の3つの源流

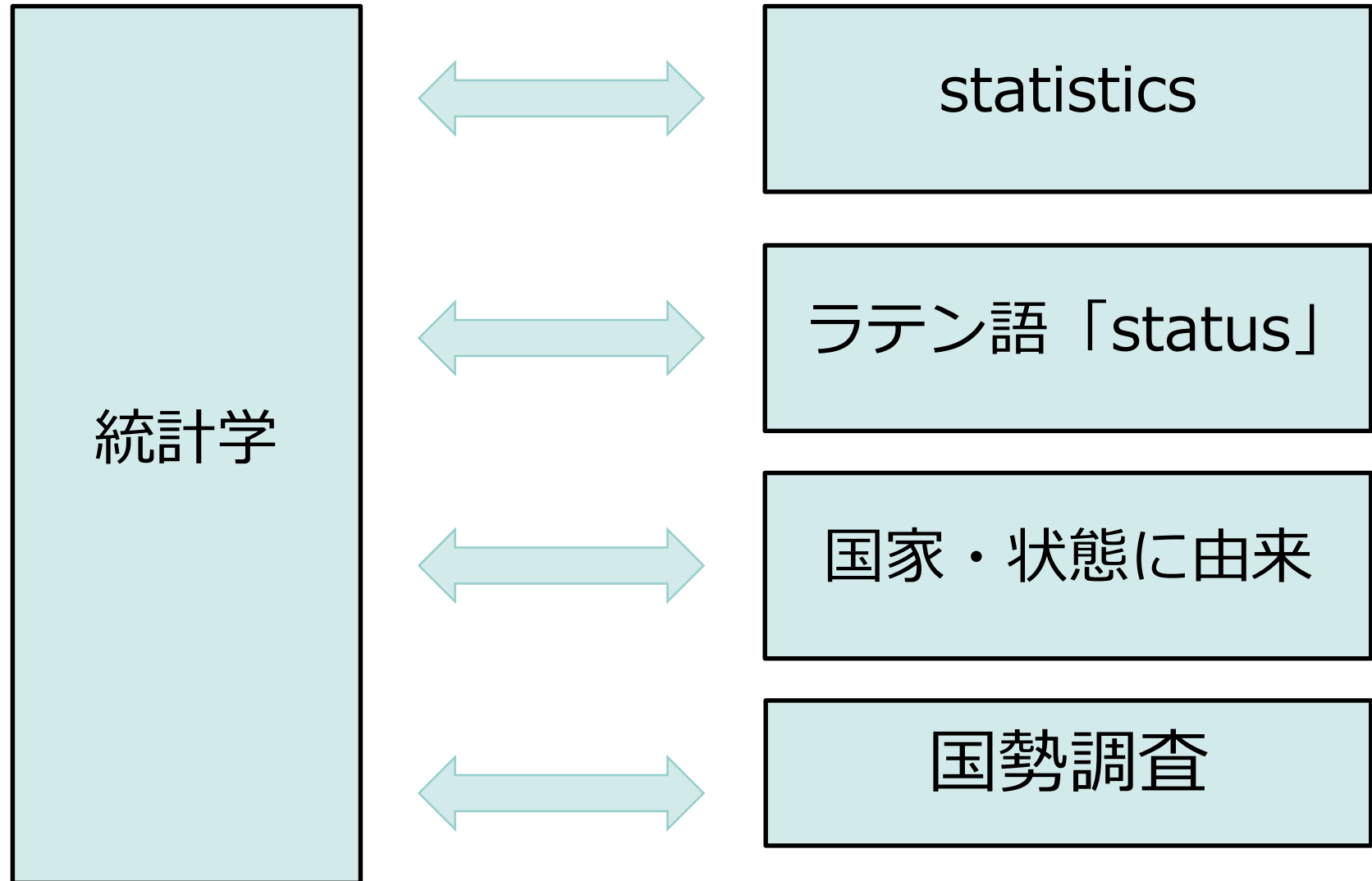
1. 国の実態を捉えるための「統計」

2. 大量の事象を捉えるための「統計」

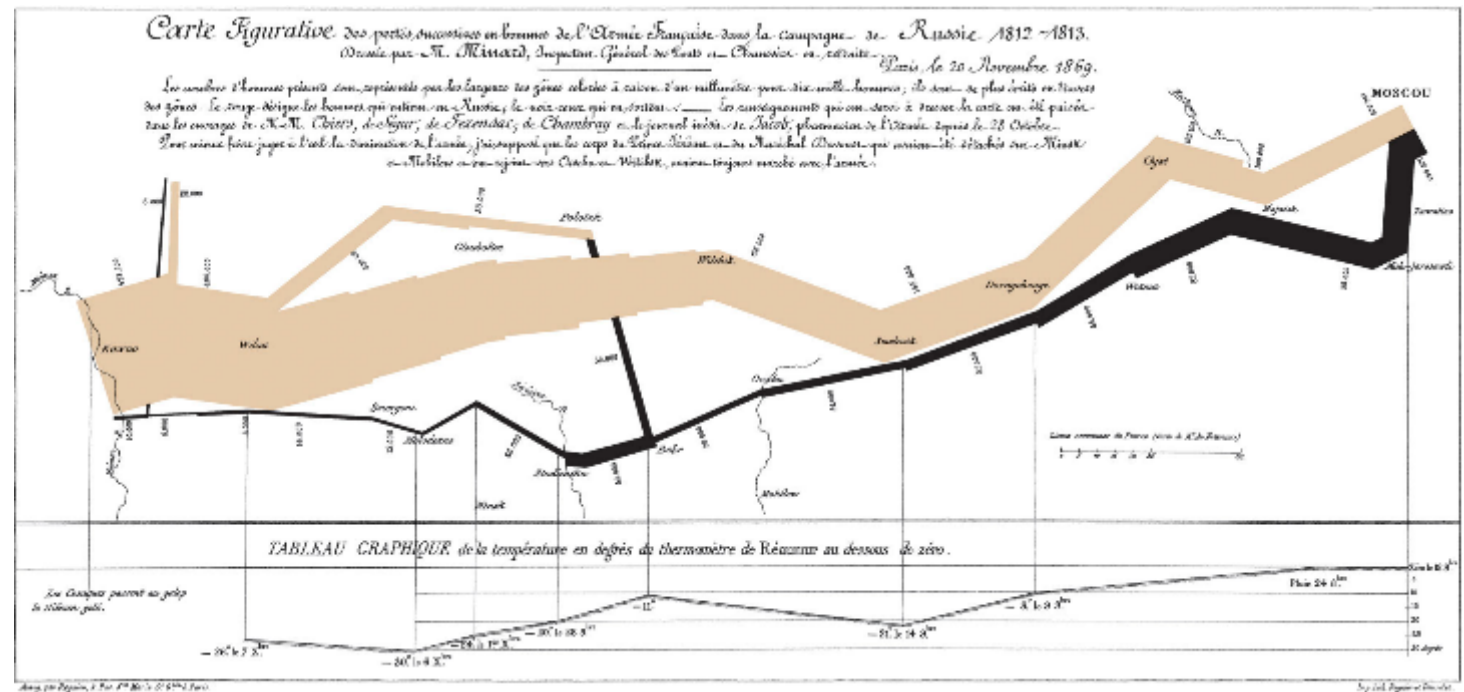
3. 確率的事象を捉えるための「統計」

現代統計学

統計学の語源



Infographic: 情報を可視化する



日本と統計学

福沢諭吉



大隈重信



福沢諭吉が説く統計学の重要性



作りすぎる



売れ残る

在庫が少なすぎる



売る機会を損失

福沢諭吉が説く統計学の重要性

経験から、季節や天気などからいくつ売れるかを予測する



雨 → 20個



晴れ → 100個

統計3級範囲が求められる場面

例：提案

提案の方法はいくつかあり、そのどの力を向上させるかによって、練習が必要となる内容が異なる。

- **資料を提出して提案を行う**

資料を解りやすくまとめる能力が必要・・・書く力（描く力）

- **資料に基づき話して提案を行う**

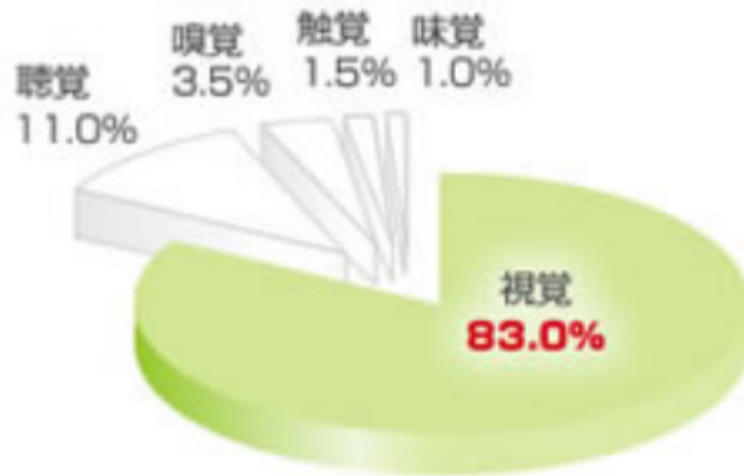
資料をうまく説明できる能力が必要・・・話す力（会話力）

人間の特性を理解する

人間の特性を理解する

出典：株式会社キテックス ビジュアルの効果

● 情報の80%は視覚によるもの



人は、五感（視覚・聴覚・嗅覚・味覚・触覚）を通して外界から情報を得て行動しており、その情報の実に80%以上を視覚から得ていることが知られています。

左記に、五感が脳神経に伝達する情報の割合をグラフに表しています。

実は、ここでも視覚効果が使われています。文章で80%と表現するよりも、図にして見るとその割合の大きさがよく分かります。

売上データの確認

月次売上表

自 平成 29年04月01日 至 平成 30年3月31日

株式会社ABC商店 9カ月経過

単位：千円

	4月	5月	6月	7月	8月	9月	10月	11月	12月	1月	2月	3月	合計
商 品 A	2,391	2,869	4,678	4,823	2,922	5,177	3,167	3,532	4,879				34,438
商 品 B	345	879	456	552	921	821	389	787	781				5,931
商 品 D	1,170	1,233	1,156	1,212	1,223	1,167	1,211	1,166	1,310				10,848
売 上 合 計	3,906	4,981	6,290	6,587	5,066	7,165	4,767	5,485	6,970				51,217

※このデータは営業がすべて販売しているとする。

▶▶▶ 演習 ▶▶▶

- ・ どんな傾向の商品なのかイメージできるか。
- ・ 仮に読み上げられて報告されたとして理解できるか。

売上データの確認

月次売上表

自 平成 29年04月01日 至 平成 30年3月31日

株式会社ABC商店

9カ月経過

単位：千円

	4月	5月	6月	7月	8月	9月	10月	11月	12月	1月	2月	3月	合計
商 品 A	2,391	2,869	4,678	4,823	2,922	5,177	3,167	3,532	4,879				34,438
商 品 B	345	879	456	552	921	821	389	787	781				5,931
商 品 D	1,170	1,233	1,156	1,212	1,223	1,167	1,211	1,166	1,310				10,848
売 上 合 計	3,906	4,981	6,290	6,587	5,066	7,165	4,767	5,485	6,970				51,217

※このデータは営業がすべて販売しているとする。

日本人はデータなどの数値を耳で聞いて理解する能力が欠如している。

そのため、上司やお客様も同様に、ヒアリングが得意ではない。

つまり、どうすれば良いのか？

提案資料や報告資料は可視化する必要がある。

情報収集の重要性

情報収集の重要性

- Scienceの基本発想
- 「現象」→「記録」→「整理」→「構造化」→「モデル化」



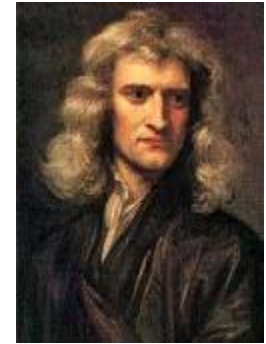
夜空には規則性がありそうだ（大昔の人）



ちゃんと記録してみよう
（ティコ・ブラーエ）



整理したら構造があった
【ケプラーの法則】
（ヨハネス・ケプラー）



運動の法則と予言
【ニュートン方程式】
（アイザック・ニュートン）

つまり、最初に必要なのは情報を
正しく収集すること。



地球は太陽の周りを回っている！
（ガリレオ・ガリレイ）

データを要約する代表値

データの種類

データ

質的データ

量的データ

名義尺度

順序尺度

間隔尺度

比率尺度

✓ 単純に区別する

✓ 順序に意味がある

✓ 差に意味がある

✓ 比に意味がある

資格の有無



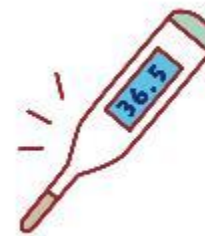
(有・無)

学校の成績



(優・良・可・不可)

温度



(38℃)

売上



(56.7億円)

データの種類

データ

質的データ

量的データ

名義尺度

順序尺度

間隔尺度

比率尺度

✓ 単純に区別する

✓ 順序に意味がある

✓ 差に意味がある

✓ 比に意味がある

- カウント
- 最頻値

- 中央値
- 分位数
(パーセンタイル)

- 算術平均
- 相関

- 前年対比
(幾何平均)
- 変動係数

ポイント

- 5つの代表値と2つの散布度
 - 代表値
 - 最大値
 - 最小値
 - 平均値
 - 中央値
 - 最頻値
 - 散布度
 - 標準偏差
 - 四分位数
- 数式はイメージをつかんで理解する
 - 統計の数式はシンプルにすると「合計して個数で割る」だけ
- 具体的な数字をExcelの関数で計算して感覚をつかむ
 - 数字に落とし込まないとイメージできない

代表値と散布度のイメージ

- まずは身近な例で感覚をつかみましょう。

#	割り勘	GC	歓送迎会
1	5,000	7,000	10,000
2	5,000	7,000	8,000
3	5,000	7,000	8,000
4	5,000	7,000	6,000
5	5,000	7,000	6,000
6	5,000	3,000	3,000
7	5,000	3,000	3,000
8	5,000	3,000	3,000
9	5,000	3,000	3,000
10	5,000	3,000	0
合計	50,000	50,000	50,000
平均値	5,000	5,000	5,000
中央値	5,000	5,000	4,500
標準偏差	0	2,000	2,933

代表値と散布度のイメージ

#	割り勘	GC	歓送迎会
1	5,000	7,000	10,000
2	5,000	7,000	8,000
3	5,000	7,000	8,000
4	5,000	7,000	6,000
5	5,000	7,000	6,000
6	5,000	3,000	3,000
7	5,000	3,000	3,000
8	5,000	3,000	3,000
9	5,000	3,000	3,000
10	5,000	3,000	0
合計	50,000	50,000	50,000
平均値	5,000	5,000	5,000
中央値	5,000	5,000	4,500
標準偏差	0	2,000	2,933

- 最大値／最小値 (maximum/minimum)
 - ・ そのデータの最も極端な値
- メリット
 - ・ データの大まかな大きさがわかる
(最大値と最小値の差は「範囲」と呼ばれる)
 - ・ 明らかにデータにそぐわない値 (異常値)
- デメリット
 - ・ データの「歪み」がある場合に適切に情報を伝えられないことがある。

EXCEL`=max(範囲)/min(範囲)`

代表値と散布度のイメージ

データを要約する代表値は？

#	割り勘	GC	歓送迎会
1	5,000	7,000	10,000
2	5,000	7,000	8,000
3	5,000	7,000	8,000
4	5,000	7,000	6,000
5	5,000	7,000	6,000
6	5,000	3,000	3,000
7	5,000	3,000	3,000
8	5,000	3,000	3,000
9	5,000	3,000	3,000
10	5,000	3,000	0
合計	50,000	50,000	50,000
平均値	5,000	5,000	5,000
中央値	5,000	5,000	4,500
標準偏差	0	2,000	2,933

	代表値	散布度
数	平均値	分散/標準偏差
順序	中央値	四分位数
その他	最頻値	範囲

平均値

#	割り勘	GC	歓送迎会
1	5,000	7,000	10,000
2	5,000	7,000	8,000
3	5,000	7,000	8,000
4	5,000	7,000	6,000
5	5,000	7,000	6,000
6	5,000	3,000	3,000
7	5,000	3,000	3,000
8	5,000	3,000	3,000
9	5,000	3,000	3,000
10	5,000	3,000	0
合計	50,000	50,000	50,000
平均値	5,000	5,000	5,000
中央値	5,000	5,000	4,500
標準偏差	0	2,000	2,933

EXCEL

=average(範囲)

=sum(範囲)/count(範囲)

- 平均値 (mean)
 - ・ 「平らに均 (なら) す」
 - ・ 一言で要約する場合は平均値を使う
 - ・ 記号「 μ (ミュー)」で表される。
- メリット
 - ・ 知名度・わかりやすさ
- デメリット
 - ・ データの「歪み」がある場合に適切に情報を伝えられないことがある。

$$\mu = \frac{1}{N} \sum_{i=1}^N x_i$$

中身

個数で割る

合計する

中央値

#	割り勘	GC	歓送迎会
1	5,000	7,000	10,000
2	5,000	7,000	8,000
3	5,000	7,000	8,000
4	5,000	7,000	6,000
5	5,000	7,000	6,000
6	5,000	3,000	3,000
7	5,000	3,000	3,000
8	5,000	3,000	3,000
9	5,000	3,000	3,000
10	5,000	3,000	0
合計	50,000	50,000	50,000
平均値	5,000	5,000	5,000
中央値	5,000	5,000	4,500
標準偏差	0	2,000	2,933

- 中央値 (median)
 - ・ 集団の中央に位置する値
 - ・ 大きい (小さい) 順に並べて中央に来る値
- メリット
 - ・ データの「歪み」がある場合にも結果が安定する。
- デメリット
 - ・ 計算が面倒 (並べ替え・場合分け)
 - ・ 知名度が低い

EXCEL =median (範囲, 戻り値)

最頻値

#	割り勘	GC	歓送迎会
1	5,000	7,000	10,000
2	5,000	7,000	8,000
3	5,000	7,000	8,000
4	5,000	7,000	6,000
5	5,000	7,000	6,000
6	5,000	3,000	3,000
7	5,000	3,000	3,000
8	5,000	3,000	3,000
9	5,000	3,000	3,000
10	5,000	3,000	0
合計	50,000	50,000	50,000
平均値	5,000	5,000	5,000
中央値	5,000	5,000	4,500
標準偏差	0	2,000	2,933

- 最頻値 (mode)
 - ・ データの中で、最も出現する頻度が高い値 (階級値)
- メリット
 - ・ ボリュームゾーンがわかる
 - ・ 最も出現可能性が高い値がわかる
- デメリット
 - ・ 全体の傾向が掴めない

EXCEL

=mode(範囲)

代表値を使いこなすコツ

- 平均値・中央値・最頻値の特徴を把握した上で活用する。
- ポイント
 - 偏り・歪みがない場合（自然現象など） 平均値≒中央値≒最頻値
 - 偏り・歪みがある場合（社会現象など） 平均値≠中央値≠最頻値

※特に、ビジネスデータ（マネー・情報）は格差が発生しやすいため、偏り・歪みが強く出る傾向にある。

代表値	メリット	デメリット
平均値	<ul style="list-style-type: none">計算しやすく高認知度推測統計の基礎情報	<ul style="list-style-type: none">外れ値の影響を受けやすい
中央値	<ul style="list-style-type: none">外れ値やノイズの影響に対して安定性がある	<ul style="list-style-type: none">並べ替え操作が必要
最頻値	<ul style="list-style-type: none">カテゴリーデータに対しても求められる	<ul style="list-style-type: none">階級の設定に依存する

標準偏差

#	割り勘	GC	歓送迎会
1	5,000	7,000	10,000
2	5,000	7,000	8,000
3	5,000	7,000	8,000
4	5,000	7,000	6,000
5	5,000	7,000	6,000
6	5,000	3,000	3,000
7	5,000	3,000	3,000
8	5,000	3,000	3,000
9	5,000	3,000	3,000
10	5,000	3,000	0
合計	50,000	50,000	50,000
平均値	5,000	5,000	5,000
中央値	5,000	5,000	4,500
標準偏差	0	2,000	2,933

EXCEL

=stdev.p(範囲)

● 標準偏差 (standard deviation)

- 偏差は平均値からのズレ
- 偏差の平均はゼロになるため、2乗したものの平均を取る。
- 平方根をとることで、2乗の効果を取り除く。

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2}$$

図解: 標準偏差の計算過程

- 個数で割る**: $\frac{1}{N}$ の部分
- 合計する**: $\sum_{i=1}^N$ の部分
- 中身**: $(x_i - \mu)^2$ の部分
- 2乗**: $(x_i - \mu)^2$ の 2 の部分
- 平方根**: $\sqrt{\quad}$ の部分
- 相殺**: 2乗と平方根が相殺する関係を示す

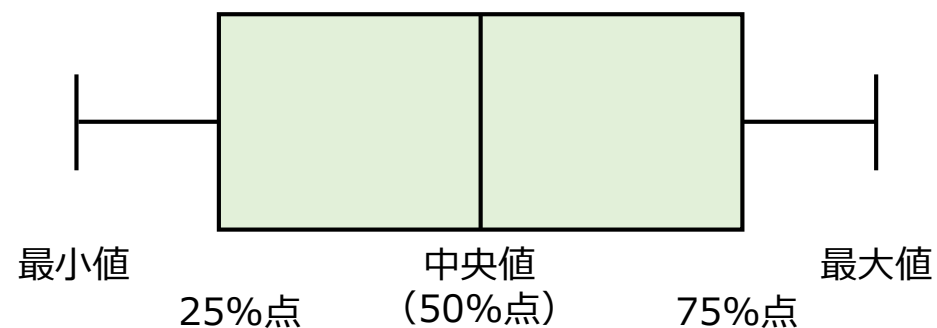
四分位数

#	割り勘	GC	歓送迎会
1	5,000	7,000	10,000
2	5,000	7,000	8,000
3	5,000	7,000	8,000
4	5,000	7,000	6,000
5	5,000	7,000	6,000
6	5,000	3,000	3,000
7	5,000	3,000	3,000
8	5,000	3,000	3,000
9	5,000	3,000	3,000
10	5,000	3,000	0
合計	50,000	50,000	50,000
平均値	5,000	5,000	5,000
中央値	5,000	5,000	4,500
標準偏差	0	2,000	2,933

EXCEL

=quartile(範囲, 戻り値)

- 四分位数 (quattile)
 - ・ データを小さい順に並べて、四つに分割する
- メリット
 - ・ データがどこに集中しているのかがわかる
 - ・ データのばらつきが把握できる
- デメリット
 - ・それほど一般的に知られていない



戻り値 : 1…最小値、2…25%点、3…中央値、4…75%点、5…最大値

データをまとめる時に重要なこと

- よくある年次の売上表も前述の平均値や標準偏差などを表に記載しておく
と理解がしやすい。
- 提案資料も少しの工夫で説明しやすくアレンジできる。



データを一言で説明する

月次売上表

自 平成 29年04月01日 至 平成 30年3月31日

株式会社ABC商店

9カ月経過

単位：千円

	4月	5月	6月	7月	8月	9月	10月	11月	12月	1月	2月	3月	合計	月平均当期	中央値	標準偏差	最小値	最大値
商品A	2,391	2,869	4,678	4,823	2,922	5,177	3,167	3,532	4,879				34,438	3,826	3,532	998	2,391	5,177
商品B	345	879	456	552	921	821	389	787	781				5,931	659	781	210	345	921
商品D	1,170	1,233	1,156	1,212	1,223	1,167	1,211	1,166	1,310				10,848	1,205	1,211	46	1,156	1,310
売上合計	3,906	4,981	6,290	6,587	5,066	7,165	4,767	5,485	6,970				51,217	5,691	5,485	1,051	3,906	7,165

○ 初めから「平均値・中央値・標準偏差・最小値・最大値」を記載しておく、データの間違いや傾向を把握しやすく、商品A、B、Cについて要約して一言で説明しやすい。

データを一言で説明する

月次売上表

自 平成 29年04月01日 至 平成 30年3月31日

株式会社ABC商店

9カ月経過

単位：千円

	4月	5月	6月	7月	8月	9月	10月	11月	12月	1月	2月	3月	合計	月平均当期	中央値	標準偏差	最小値	最大値
商品A	2,391	2,869	4,678	4,823	2,922	5,177	3,167	3,532	4,879				34,438	3,826	3,532	998	2,391	5,177
商品B	345	879	456	552	921	821	389	787	781				5,931	659	781	210	345	921
商品D	1,170	1,233	1,156	1,212	1,223	1,167	1,211	1,166	1,310				10,848	1,205	1,211	46	1,156	1,310
計	3,906	4,981	6,290	6,587	5,066	7,165	4,767	5,485	6,970				51,217	5,691	5,485	1,051	3,906	7,165

データをまとめる時に重要なこと

- よくある年次の売上表も前述の平均値や標準偏差などを表に記載しておく
と理解がしやすい。
- 提案資料も少しの工夫で説明しやすくアレンジできる。



データを一言で説明する

月次売上表

自 平成 29年04月01日 至 平成 30年3月31日

株式会社ABC商店

9カ月経過

単位：千円

	4月	5月	6月	7月	8月	9月	10月	11月	12月	1月	2月	3月	合計	月平均当期	中央値	標準偏差	最小値	最大値
商品A	2,391	2,869	4,678	4,823	2,922	5,177	3,167	3,532	4,879				34,438	3,826	3,532	998	2,391	5,177
商品B	345	879	456	552	921	821	389	787	781				5,931	659	781	210	345	921
商品D	1,170	1,233	1,156	1,212	1,223	1,167	1,211	1,166	1,310				10,848	1,205	1,211	46	1,156	1,310
計	3,906	4,981	6,290	6,587	5,066	7,165	4,767	5,485	6,970				51,217	5,691	5,485	1,051	3,906	7,165

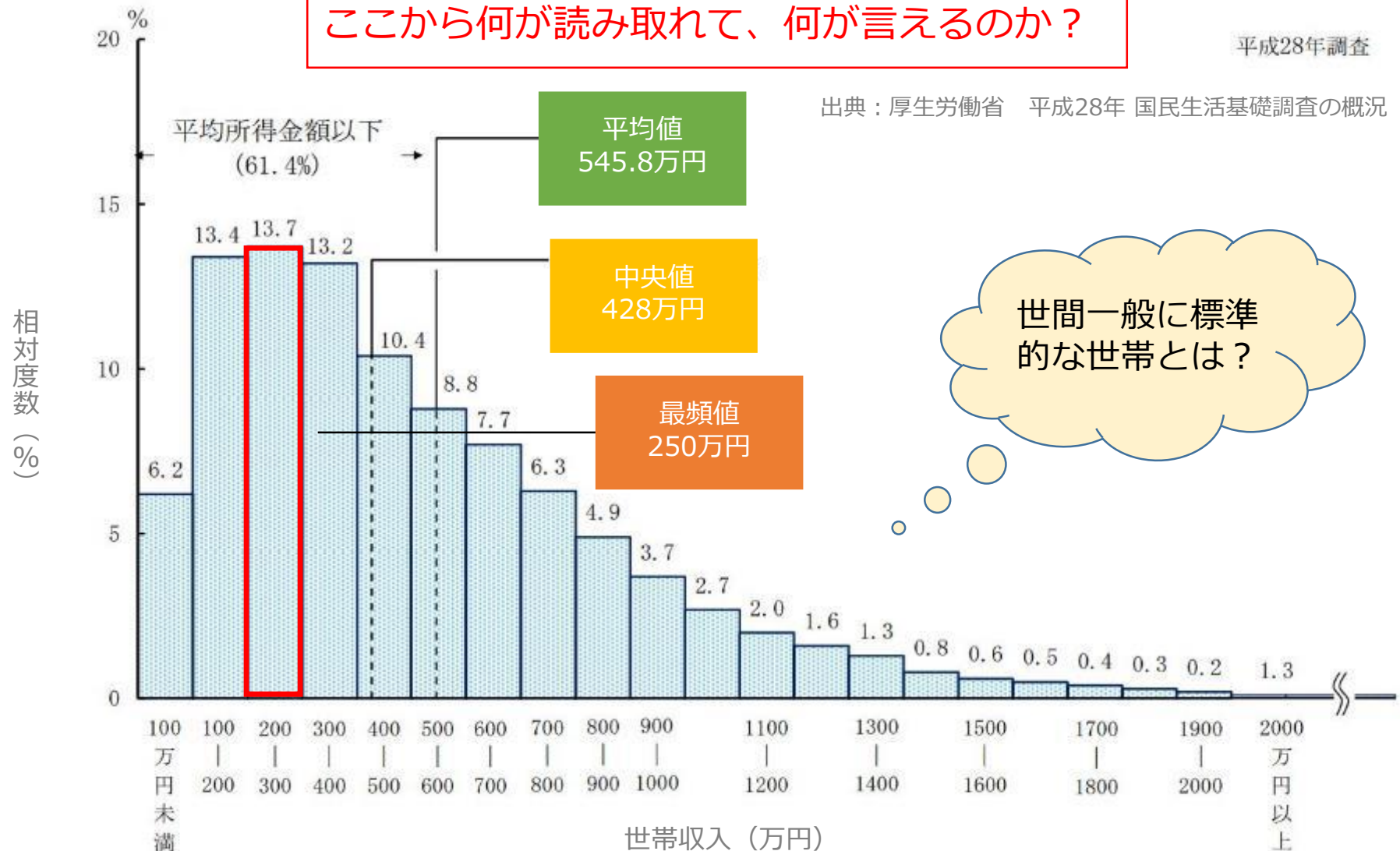
○ 初めから「平均値・中央値・標準偏差・最小値・最大値」を記載しておく、データの間違いや傾向を把握しやすく、商品A、B、Cについて要約して一言で説明しやすい。

代表値に差が表れる分布

ここから何が読み取れて、何が言えるのか？

平成28年調査

出典：厚生労働省 平成28年 国民生活基礎調査の概況



統計検定 3級試験対策 ～概要～

統計検定3級 概要

- 試験日時
 - ✓毎年6月,11月に実施
- 試験概要
 - ✓試験時間60分で30問
- 合格基準
 - ✓70%で合格（公式発表）
 - ✓目安となる合格ラインは21問



統計検定
Japan Statistical Society Certificate

統計検定 3 級 出題内容

統計検定3級出題範囲表

大項目	小項目	ねらい	項目(学習しておくべき用語)
データの種類	データの基礎知識	データのタイプの違いを理解し、それぞれのデータに適した処理法を理解する。	量的データ、質的データ、名義尺度、順序尺度、間隔尺度、比率尺度、連続尺度
標本調査	母集団と標本	標本調査の意味と必要性を理解し、標本の抽出方法や推定方法について説明することができる。	母集団、標本、全数調査、無作為抽出、標本の大きさ、乱数表、国勢調査
実験調査	実験の基本的な考え方	実験調査の意味と必要性を理解し、実験の基本的な考え方について、説明することができる。	処理群と対照群
統計グラフ	1変数の基本的なグラフの見方・読み方	基本的な1変数の統計グラフを適切に解釈したり、自ら書いたりすることができる。	棒グラフ、折れ線グラフ、円グラフ、帯グラフ
	2変数の基本的なグラフの見方	基本的な2変数の統計グラフを適切に解釈したり、自ら書いたりすることができる。	モザイク図、箱ひげ図、散布図、複合グラフ
データの集計	1 変数データ	1変数のデータを適切に集計表に記述すること、また集計表から適切に情報を読み取り、説明することができる。	度数分布表、度数、相対度数、累積度数、累積相対度数、階級、階級値、度数分布表からの統計量の求め方
	2 変数データ	2変数のデータを適切にクロス集計表に記述すること、また集計表から適切に情報を読み取り、説明することができる。	クロス集計表(2 元の度数分布表)
データの代表値	代表値とその利用法	数値を用いてデータの中心的位置を表現すること、またそれらを用いて適切にデータの特徴を説明することができる。	平均値、中央値、最頻値
データの散らばり	量的な1変数の散らばりの指標	データの散らばりを、指標を用いて把握し、説明することができる。	最小値、最大値、範囲、四分位数、四分位範囲(四分位偏差)、分散、標準偏差、偏差値、変動係数
	量的な2変数の散らばりの指標	量的な2つの変数の散らばりを指標から把握し、説明することができる。	共分散、相関係数
	散らばりのグラフ表現	データの散らばりをグラフ表現することを通して、散らばりの特徴を把握したり、グループ間の比較を行ったりすることができる。外れた値の処理を考える。	ヒストグラム(柱状グラフ)、累積相対度数グラフ、箱ひげ図、散布図(相関図)、はずれ値
確率	確率の基礎	確率の意味や基本的な法則を理解し、さまざまな事象の確率を求めたり、確率を用いて考察することができる。	独立な試行、条件付き確率
時系列データ	時系列データの基本的な見方	時系列情報を持つデータをグラフや指標を用いて適切に表現し、それらの情報を適切に読み取ることができる。	時系列グラフ、指標化

統計3級出題内容（2018年11月実施）

大問	小問	出題範囲	大問	小問	出題範囲
問1		データの種類	問10	[2]	データの可視化
問2		集合と確率	問11		データの集計
問3		集合と確率	問12		データの集計
問4		データの可視化	問13	[1]	相関係数／共分散
問5	[1]	データの可視化		[2]	相関係数／共分散
	[2]	データの可視化	問14		相関係数／共分散
	[3]	データの可視化	問15		相関係数／共分散
問6		データの可視化	問16	[1]	データの可視化
問7	[1]	要約統計量		[2]	データの可視化
	[2]	要約統計量		[3]	相関係数／共分散
問8	[1]	要約統計量	問17		要約統計量
	[2]	要約統計量	問18	[1]	データの可視化
問9	[1]	データの可視化		[2]	データの可視化
	[2]	データの可視化	問19	[1]	標本調査
問10	[1]	データの可視化		[2]	標本調査

問題演習



統計3級出題内容（2018年11月実施）

大問	小問	出題範囲	大問	小問	出題範囲
問1		データの種類	問10	[2]	データの可視化
問2		集合と確率	問11		データの集計
問3		集合と確率	問12		データの集計
問4		データの可視化	問13	[1]	相関係数／共分散
問5	[1]	データの可視化		[2]	相関係数／共分散
	[2]	データの可視化	問14		相関係数／共分散
	[3]	データの可視化	問15		相関係数／共分散
問6		データの可視化	問16	[1]	データの可視化
問7	[1]	要約統計量		[2]	データの可視化
	[2]	要約統計量		[3]	相関係数／共分散
問8	[1]	要約統計量	問17		要約統計量
	[2]	要約統計量	問18	[1]	データの可視化
問9	[1]	データの可視化		[2]	データの可視化
	[2]	データの可視化	問19	[1]	標本調査
問10	[1]	データの可視化		[2]	標本調査