

情報Ⅱ『情報とデータサイエンス』に おける指導の在り方

情報Ⅱ第3領域で扱うデータサイエンス

日出学園中学校・高等学校
武善 紀之

<https://www.hinode.ed.jp/share/takeyoshi/>



自己紹介

自己紹介

名前 武善紀之(たけよし のりゆき)

所属 私立 日出学園中学校・高等学校 教諭
所在地:千葉県市川市 ※幼稚園・小学校併設

教科 情報科 12年目(+数学,公民,技術)

現在 高3担任・進路指導部・パソコン部顧問

年齢 34歳

出身 筑波大学情報学群情報メディア創成学類

好きなもの ペンギン



<https://www.hinode.ed.jp/share/takeyoshi/>

「情報科」関連の主な兼職

・教科書執筆

- ・情報科「新編情報Ⅰ」「情報Ⅰ Step Forward!」「情報Ⅱ」(東京書籍)

・予備校

- ・「ベーシックレベル情報Ⅰ」講師(スタディサプリ)

・教育番組

- ・NHK高校講座「情報Ⅰ」監修講師(NHK)

・教員研修

- ・GIGA スクールにおける学びの充実「高等学校情報教員指導力向上事業」(文部科学省)
- ・GIGA スクール「情報Ⅱ」高等学校情報科等強化によるデジタル人材の供給体制整備支援事業(文部科学省)

スタディサプリ



NHK 高校講座



武善 紀之
たけよし のりゆき
日出国中学校・高等学校教諭







ペンギンが好きすぎて……



教員南極派遣プログラム

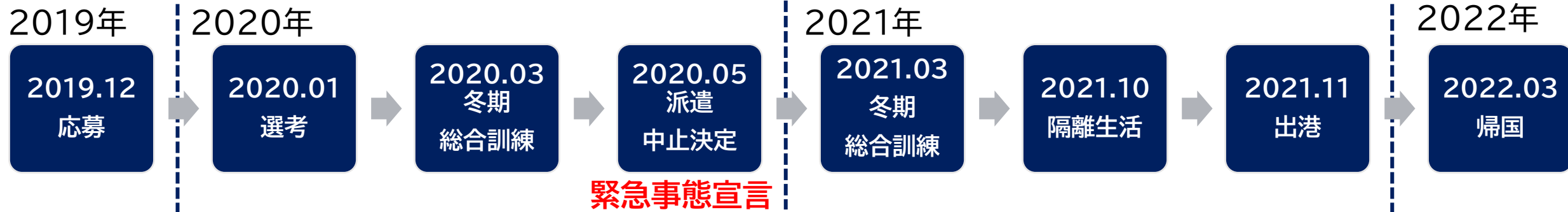
- 国立極地研究所が主催(2009年～)
- 極地の科学や観測に興味を持つ
現職教員を南極昭和基地に派遣
※日本南極地域観測隊(JARE)に同行
- 派遣予定人数 2名
(例年、9月下旬～12月上旬が募集期間)
- 衛星回線を利用して、
現地から「南極授業」を行う

(参考)2023年度教員南極派遣プログラム派遣教員募集ページ
<https://www.nipr.ac.jp/antarctic/info/2023teacher/>

情報科教員派遣は
初！



スケジュール



選考

- 1次:書類(授業案2本&帰国後の活動計画提出)
2次:隊長経験者と面接

計画1 南極×プログラミング(設営部門に着目)

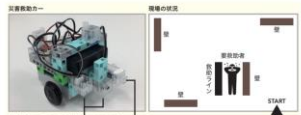
【授業の概要】

プログラミングで、「南極の〇〇を解決せよ！」

あなたは南極観測に参加した、設営隊員・観測隊員の一人。現地で動作する「あるシステム」の開発に携わっていたが、現地で使用した結果、そのままでは満足に動作しないことが判明した。プログラミングの環境、各種センサ類は揃っている。知恵と技術で、円滑な観測環境の設置を実現せよ！

【授業のイメージ】

ハードウェアも組み合わせたプログラミングの問題解決学習は情報科でも主流であり、私自身3年前から継続して研究を行っている。基本的なプログラミング手法、制御装置について学習した後、過去には「災害救助力一を制御しよう」というタイトルで、問題解決学習を実施した(50分設定)。同内容は、(株)東京書籍の「まづはここから プログラミング実践事例集2」に掲載されている。



計画2. 南極×データサイエンス(観測部門に着目)

【授業の概要】

データ分析で、「南極の〇〇を解きあかせ！」

あなたは南極観測に参加した、生物圏観測隊員の一人。現地で生物圏の調査を行うために、アデリーペンギンのデータロガーから、あるデータを入手した。データを分析し、ペンギンの意外な生態系を解き明かせ！

【授業のイメージ】

「バイオリギングを用いたデータ解析」を生徒自身が体験する授業を実施する。私自身が強く感銘を受けたのは例えば以下の事例であった。地上で観察するだけでは確認できない事象について、動物たちに小型計測装置を取り付け、データ解析することで、動物たちの本当の生活実態が浮かび上がる。以前より、IoTやビックデータの単元を授業で扱う際には、必ず以下の事例を紹介していた。

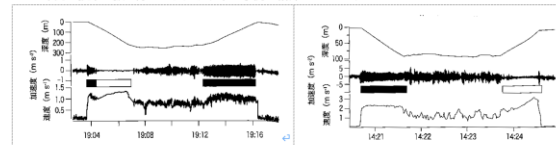


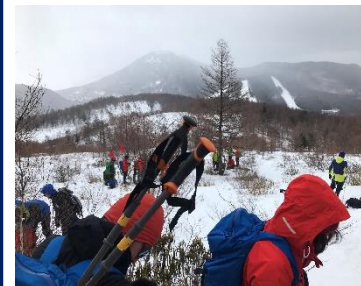
図1 バイカルアラザラシから得られた潜水中の深度・加速度・速度。足ヒレを動かしていた時間を黒横棒、動かしていなかった時間を白横棒(5より)。

図2 ジェンツーペンギンから得られた潜水中の深度・加速度・速度。翼を動かしていた時間を黒横棒、動かしていなかった時間を白横棒(5より)。

多くは理科の先生(2009年～),
情報科の教員派遣例は“初”!

冬期総合訓練 場所:長野県東御市 湯の丸高原一帯

雪上歩行訓練



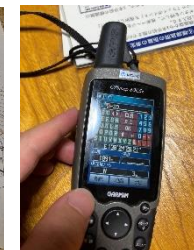
クレバス脱出訓練



貸与装備



ルート工作訓練



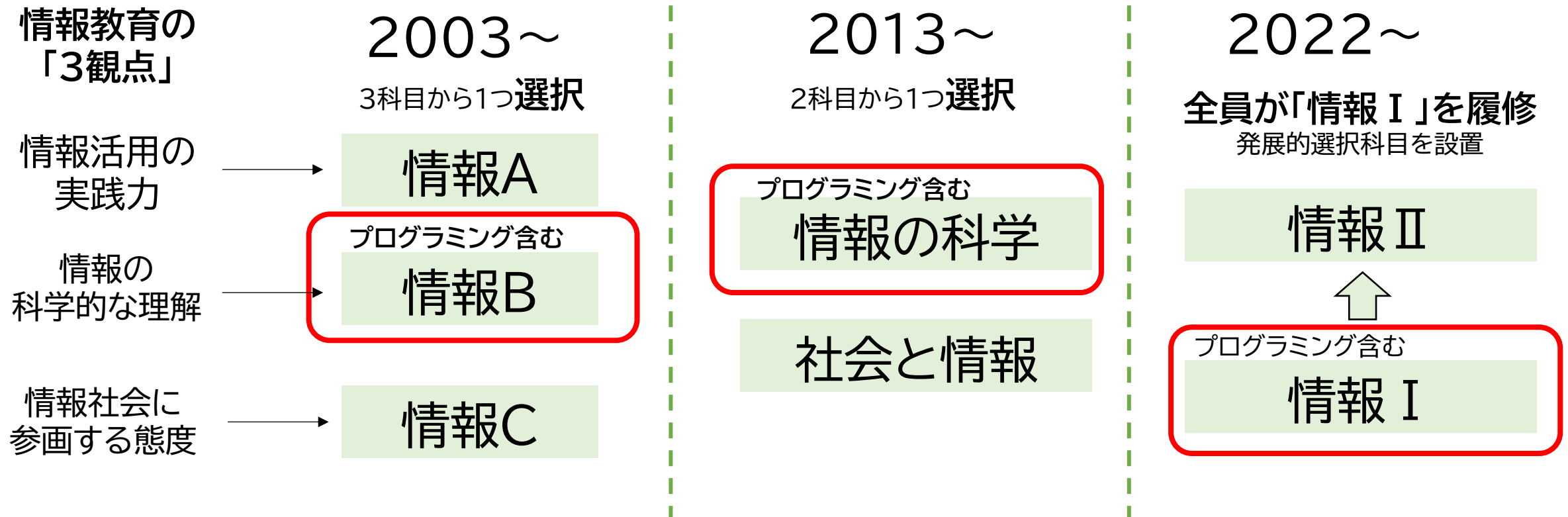
雪山テント泊



Agenda.

- 1 情報Ⅱのデータサイエンス (情報Ⅰと比較しながら)
- 2 情報Ⅱの授業演習
- 3 情報Ⅰの逆向き設計
- 4 データサイエンスを扱う価値

情報Ⅰと情報Ⅱに関するざっとしたおさらい



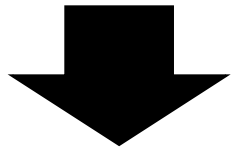
×新教科「情報Ⅰ」では扱う内容が大幅に増加！

○「プログラミング」や「シミュレーション」については、
今まで「選択」であった内容が「必修」になった。

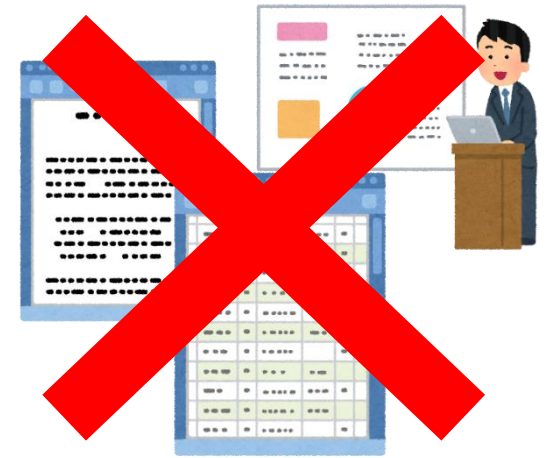
新学習指導要領「情報Ⅰ」

「情報Ⅰ」のコンセプト

1. コンピュータを活用して、**問題解決**ができるようになること
2. **情報科学(Computer Science)**をきちんと学ぶこと

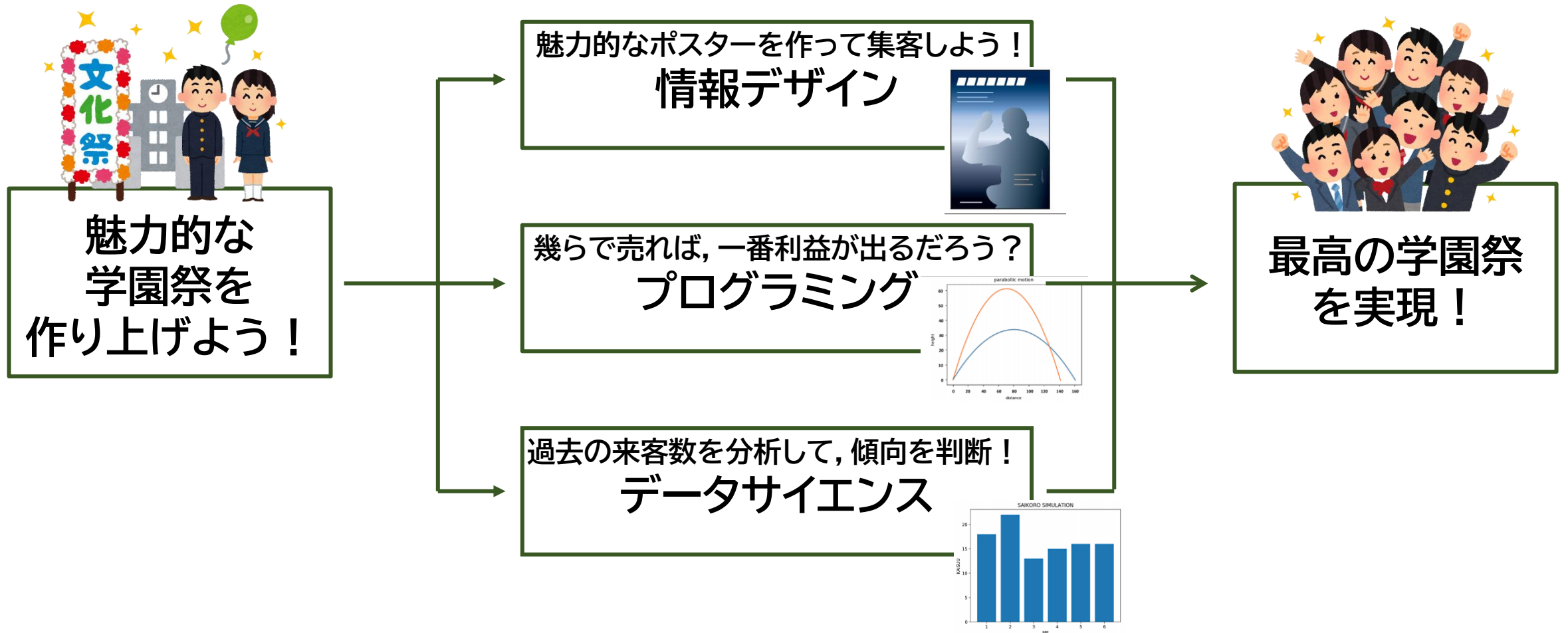


- ◆第1章 情報社会の問題解決
- ♥第2章 コミュニケーションと情報デザイン
- ♣第3章 コンピュータとプログラミング
- ♠第4章 情報通信ネットワークとデータの活用



コンピュータを活用して問題解決ができるようになる

「情報デザイン」「データサイエンス」「プログラミング」は、
問題解決の三種の神器

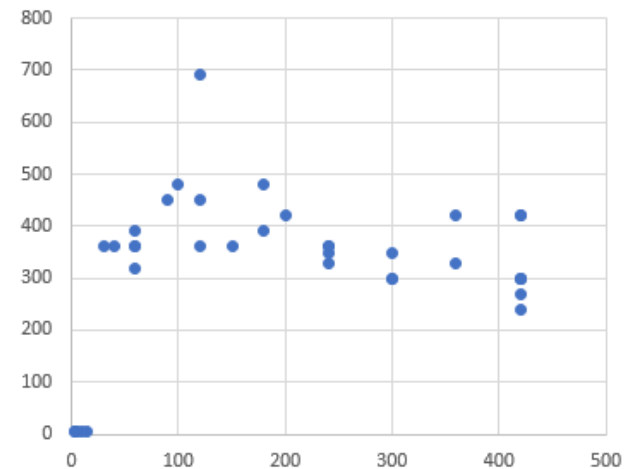


情報 I で扱われるデータサイエンスのイメージ

情報 I の学習内容は、
大雑把にまとめると**2つ**

1. **自分達でアンケート集計や分析を行う内容**
2. **データベースや情報システムに関する内容**

1日の平日のスマ	1日の平日の平均
240	350
14	5
240	330
360	330
60	360
120	690



キャラクター 涼しい 幅広い ぼつちやり
高い 連想 ペンギンハイウェイ 速い
すいか マスコットキャラクター 飛べる 滑る
動物園 エイプリル 愛着 とる 王道 水族館 北極 集まる
生き抜く 氷 顧問 武善 ピンギー ペンギン
うまい 泳ぐ とべる suica 飛べない鳥 過酷 育てる
歩く 上 群れ かりる yochiyochi エヴァ
可愛い 住む 群れ 南極 トリ 生息 かわいい
ペンギンの問題 しろくまカフェ ペンギンさん
持つ 寒い スイカ 先生 コラ画像 かつこい

sAccess: データベース実習支援ツール

DB選択に戻る リンク

ゲストさん ようこそ!

操作コマンドを追加しよう

ここにコマンドを入力

追加

操作コマンド一覧

表示 売上データ

チェックしたコマンドを

一つ上へ 一つ下へ 削除

テーブル確認&更新【コンビニ】

・売上データ (158件)

・商品データ (30件)

操作結果ダウンロード(CSV)

テーブル&命令列ダウンロード(.saccess)

売上データ (158件)

	商品コード	売上日	曜日	時間帯	性別	年齢層
1	G6148	4/1	日	朝	男	若者
2	J0940	4/1	日	朝	女	若者
3	S6356	4/1	日	朝	男	成年
4	S4436	4/1	日	昼	女	成年
5	G3944	4/1	日	昼	男	子ども
6	T0344	4/1	日	昼	女	子ども
7	S6356	4/1	日	夕方	男	若者
8	J0589	4/1	日	夕方	女	熟年
9	S4777	4/1	日	夜	男	熟年
10	J0589	4/1	日	夜	女	若者
11	T6962	4/1	日	深夜	男	成年
12	S6356	4/2	月	朝	女	熟年

<=[表示]=

これを受けて情報Ⅱでは…？

データサイエンスの学習内容は、
大雑把にまとめると**3つ**

1. 自分達でアンケート集計や分析を行う内容 →主に、「情報Ⅰ」
2. データベースや情報システムに関する内容 →両方
3. **大規模データの分析,可視化に関する内容 →主に「情報Ⅱ」**

ともかく大規模！
10次元以上とか当たり前。

	01_睡眠	02_身の回りの用事	03_食事	04_通勤・通学	05_仕事	06_学業	07_家事	08_買い物	11_移動(通勤・通学を除く)	12_テレビ・ラジオ・新聞・雑誌	13_休養・くつろぎ	14_学習・自己啓発・訓練(学業以外)	15_趣味・娯楽	16_スポーツ
01_北海道	467	72	74	62	41	355	5	3	19	25	166	42	63	18
02_青森県	469	71	90	53	34	368	0	0	13	79	128	19	74	30
03_岩手県	458	69	76	53	47	453	5	3	8	31	105	42	59	14
04_宮城県	448	75	80	72	44	373	3	3	7	16	93	29	110	53
05_秋田県	467	61	84	45	29	373	4	6	13	19	136	32	114	34
06_山形県	477	74	83	52	43	377	4	5	6	40	135	38	58	29
07_福島県	436	60	87	64	41	438	7	2	15	23	146	36	40	25
08_茨城県	456	88	76	71	31	387	6	6	18	24	133	41	48	35
09_栃木県	472	66	82	69	46	394	13	12	9	24	131	13	72	20
10_群馬県	457	74	78	69	41	398	7	7	12	31	125	36	67	22
11_埼玉県	439	83	91	80	31	401	5	3	12	28	134	55	41	13
12_千葉県	437	71	85	85	70	367	8	3	12	15	170	54	33	17
13_東京都	451	84	86	80	33	387	5	7	11	16	124	55	56	36
14_神奈川	452	89	86	77	32	399	2	8	14	39	108	23	78	9
15_新潟県	475	71	74	50	13	381	7	11	8	49	143	49	62	21
16_富山県	475	66	75	66	68	376	3	4	10	35	125	28	64	15
17_石川県	460	69	78	51	48	395	3	1	16	16	135	31	79	43
18_福井県	445	87	75	56	36	450	3	4	10	12	126	54	45	25
19_山梨県	453	75	78	62	28	429	5	1	16	30	94	60	66	14
20_長野県	463	62	91	72	43	402	10	1	6	26	150	30	53	17
21_岐阜県	458	73	78	69	38	418	13	6	9	45	101	36	57	23
22_静岡県	452	82	81	62	49	390	8	3	11	21	136	41	54	23
23_愛知県	459	93	79	71	72	374	4	1	12	25	123	44	51	13
24_三重県	437	66	75	81	87	358	3	2	6	21	171	42	41	23
25_滋賀県	462	81	83	60	47	398	5	4	6	12	144	46	47	27
26_京都府	478	65	83	53	23	360	17	10	14	27	159	34	53	15
27_大阪府	430	88	84	70	65	366	8	6	16	23	147	56	51	22
28_兵庫県	465	79	87	69	45	351	3	6	12	33	161	40	48	19
29_奈良県	446	72	86	107	10	416	7	15	9	41	75	36	100	13
30_和歌山	467	58	81	71	31	403	7	6	14	37	139	31	45	23
31_鳥取県	464	69	72	55	55	408	9	4	13	37	138	38	27	31
32_島根県	449	68	75	66	47	440	1	1	8	14	152	51	37	19
33_岡山県	424	88	82	77	40	405	4	2	21	14	135	49	40	11
34_広島県	450	84	80	67	27	433	18	6	7	20	132	28	71	17
35_山口県	432	70	81	67	63	394	7	5	20	25	133	34	63	32
36_徳島県	451	76	77	52	65	401	3	3	8	31	111	61	64	18
37_香川県	450	61	70	67	33	388	5	6	11	20	110	60	60	10

指導要領の言葉を比較してみる

情報Ⅰ

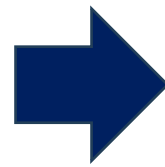
(1) 情報社会の問題解決
(2) コミュニケーションと情報デザイン
(3) コンピュータとプログラミング
(4) 情報通信ネットワークとデータの活用

情報Ⅱ

(1) 情報社会の進展と情報技術
(2) コミュニケーションとコンテンツ
(3) 情報とデータサイエンス
(4) 情報システムとプログラミング
(5) 情報と情報技術を活用した問題の発見・解決の探究

図版: 高等学校情報科「情報Ⅰ」教員研修用教材(文部科学省)

情報Ⅰ(必修) ※学習指導要領解説より
質的データと量的データ, データの整理(データクリーニングなど), 欠損値や外れ値の扱い方, データの様々な可視化, テキストマイニングオープンデータの活用, 散布図行列, 相関と交絡因子, 単回帰分析, 統計的仮説検定



情報Ⅱ(選択科目) ※学習指導要領解説より
選択バイアス, 情報バイアス, **機械学習**, **訓練データ**, **過剰適合**, **重回帰分析**, **モデルの調整**, **主成分分析**, **クラスタリング**, **文字認識**, **画像認識**, **人工知能**

とてもじゃないけど、できない？

機械学習をまとめると

機械学習

機械が、データからパターンや規則を学習し、予測や判断を行う技術。
AI(人工知能)を実現する技術の1つ。

正解データがある

正解データがない

教師あり学習

数値の予測

回帰

Regression

- ・線形回帰

データの分類

クラス分類

Classification

- ・ロジスティック回帰
- ・決定木

教師なし学習

主要な特徴を
残す・取り出す

次元削減

Dimensionality
Reduction

- ・主成分分析

データの
グループ化

クラスタリング

Clustering

- ・k平均法

やっぱり、とてもじゃないけど、できない？

実は「情報Ⅰ」より「情報Ⅱ」の方が
面白い／簡単？

ここまで無理だと諦めがつく

ともかく大規模！
10次元以上とか当たり前。

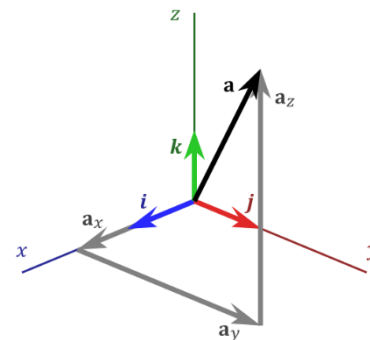
	01_睡眠	02_身の回りの用事	03_食事	04_通勤・通学	05_仕事	06_学業	07_家事	10_買い物	11_移動 (通勤・通学を除く)	12_テレビ・ラジオ・新聞・雑誌	13_休養・くつろぎ	14_学習・自己啓発・訓練(学業以外)	15_趣味・娯楽	16_スポーツ
01_北海道	467	72	74	62	41	355	5	3	19	25	166	42	63	18
02_青森県	469	71	90	53	34	368	0	0	13	79	128	19	74	30
03_岩手県	458	69	76	53	47	453	5	3	8	31	105	42	59	14
04_宮城県	448	75	80	72	44	373	3	3	7	16	93	29	110	53
05_秋田県	467	61	84	45	29	373	4	6	13	19	136	32	114	34
06_山形県	477	74	83	52	43	377	4	5	6	40	135	38	58	29
07_福島県	436	60	87	64	41	438	7	2	15	23	146	36	40	25
08_茨城県	456	88	76	71	31	387	6	6	18	24	133	41	48	35
09_栃木県	472	66	82	69	46	394	13	12	9	24	131	13	72	20
10_群馬県	457	74	78	69	41	398	7	7	12	31	125	36	67	22
11_埼玉県	439	83	91	80	31	401	5	3	12	28	134	55	41	13
12_千葉県	437	71	85	85	70	367	8	3	12	15	170	54	33	17
13_東京都	451	84	86	80	33	387	5	7	11	16	124	55	56	36
14_神奈川	452	89	86	77	32	399	2	8	14	39	108	23	78	9
15_新潟県	475	71	74	50	13	381	7	11	8	49	143	49	62	21
16_富山県	475	66	75	66	68	376	3	4	10	35	125	28	84	15
17_石川県	460	69	78	51	48	395	3	1	16	16	135	31	79	43
18_福井県	445	87	75	56	36	450	3	4	10	12	126	54	45	25
19_山梨県	453	75	78	62	28	429	5	1	16	30	94	60	66	14
20_長野県	463	62	91	72	43	402	10	1	6	26	150	30	53	17
21_岐阜県	458	73	78	69	38	418	13	6	9	45	101	36	57	23
22_静岡県	452	82	81	62	49	390	8	3	11	21	136	41	54	23
23_愛知県	459	93	79	71	72	374	4	1	12	25	123	44	51	13
24_三重県	437	66	75	81	87	358	3	2	6	21	171	42	41	23
25_滋賀県	462	81	83	60	47	398	5	4	6	12	144	46	47	27
26_京都府	478	65	83	53	23	360	17	10	14	27	159	34	53	15
27_大阪府	430	88	84	70	65	366	8	6	16	23	147	56	51	22
28_兵庫県	465	79	87	69	45	351	3	6	12	33	161	40	48	19
29_奈良県	446	72	86	107	10	416	7	15	9	41	75	36	100	13
30_和歌山	467	58	81	71	31	403	7	6	14	37	139	31	45	23
31_鳥取県	464	69	72	55	55	408	9	4	13	37	138	38	27	31
32_島根県	449	68	75	66	47	440	1	1	8	14	152	51	37	19
33_岡山県	424	88	82	77	40	405	4	2	21	14	135	49	40	11
34_広島県	450	64	80	67	27	433	18	6	7	20	132	28	71	17
35_山口県	432	70	81	67	63	384	7	5	20	25	133	34	63	32
36_徳島県	451	76	77	52	65	401	3	3	8	31	111	61	64	18
37_香川県	450	61	76	67	29	388	5	6	11	26	116	29	62	16

人工知能(AI)やコンピュータを
活用しないと、無理！

(情報 I はギリギリ「できてしまう」範囲だった)
コンピュータに頼らざるを得ない。

※僕のイメージ

高校数学のベクトル
(3次元まで)



大学の線形代数
(n次元空間)

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

※図はWikipediaより

指導要領もよく読むと……

（情報Ⅱ学習指導要領解説）

ここでは、数学科における学習内容と関連する部分も含むが、**数学や統計学の専門的な内容に深入りすることなく、可視化やソフトウェアによる処理の結果を基に、その概念を理解するようにする。**

（実際に指導してみると……）

- 情報Ⅰより、**ソフトウェアの操作が簡単**（コマンド入力で終わり）
- 情報Ⅰより、**数学感が薄い**（うわ、苦手だった奴……がない）
- 情報Ⅰより、**扱えるデータが多様**（みんな全く違うテーマで楽しめる）

（生徒も）

- 情報Ⅱは実践的なものが多かったので**楽しい**。
- たくさんのデータをまとめることは数学などでは難しいので、それらをいっきにまとめられるのは**とても便利だと思った**。

僕がイメージしているゴール設計

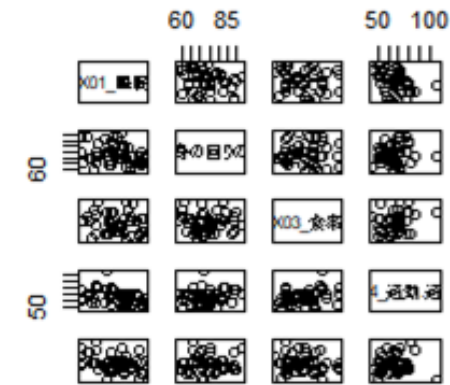
多次元のデータ

	01.睡眠	02.身の回りの	03.食事	04.通勤・通学	05.仕事	06.学業	07.家事	08.買い物	09.移動(通)	10.テレビ	11.休養・趣味・娯楽	12.その他	13.身体健康	14.学習・読書	15.社会生活	16.その他
01.睡眠	467	72	74	62	41	355	5	9	19	25	166	42	63	18		
02.身の回りの	469	71	80	53	34	386	5	9	19	25	166	42	63	18		
03.食事	459	69	76	53	47	453	5	9	19	25	166	42	63	18		
04.通勤・通学	449	75	72	44	173	3	9	19	25	166	42	63	18			
05.仕事	467	61	84	45	29	373	4	6	10	19	136	32	114	34		
06.学業	477	14	63	52	40	377	4	6	10	19	136	32	114	34		
07.家事	439	60	87	64	41	439	7	2	15	29	149	36	40	25		
08.買い物	439	60	87	64	41	439	7	2	15	29	149	36	40	25		
09.移動(通)	439	60	87	64	41	439	7	2	15	29	149	36	40	25		
10.テレビ	439	60	87	64	41	439	7	2	15	29	149	36	40	25		
11.休養・趣味・娯楽	439	60	87	64	41	439	7	2	15	29	149	36	40	25		
12.その他	439	60	87	64	41	439	7	2	15	29	149	36	40	25		
13.身体健康	439	60	87	64	41	439	7	2	15	29	149	36	40	25		
14.学習・読書	439	60	87	64	41	439	7	2	15	29	149	36	40	25		
15.社会生活	439	60	87	64	41	439	7	2	15	29	149	36	40	25		
16.その他	439	60	87	64	41	439	7	2	15	29	149	36	40	25		

情報 I

2つずつで
頑張ってみる

	01 睡眠	02 身の回りの	03 食事	04 通勤・通学	05 仕事	06 学業
01 睡眠						
02 身の回りの	-0.2308					
03 食事	-0.129	0.04927				
04 通勤・通学	-0.3563	0.19467	0.28387			
05 仕事	-0.3274	0.08675	-0.1459	0.02845		
06 学業	-0.3835	-0.1301	-0.2055	-0.0192	-0.0992	
07 家事	0.03418	-0.2337	0.04911	0.06168	-0.1948	0.05732
08 買い物	0.26421	0.02212	0.0845	0.12755	-0.4664	-0.239
09 移動(通)	-0.1549	0.14999	0.02546	-0.107	-0.1786	-0.2163
10 テレビ	0.35748	-0.1084	0.11338	-0.0222	-0.2739	-0.0922
11 休養・趣味・娯楽	0.09791	-0.1746	0.04584	-0.1438	0.2955	-0.4069
12 学習・読書	-0.1175	0.37528	-0.0702	0.03141	0.02206	-0.2363
13 趣味・娯楽	0.12724	-0.1379	0.11829	-0.052	-0.3413	-0.2828
14 スポーツ	0.0383	-0.0809	-0.0632	-0.2982	0.01111	-0.1535
20 その他	0.10732	-0.0867	-0.2132	-0.3353	-0.1576	0.01509



相関行列

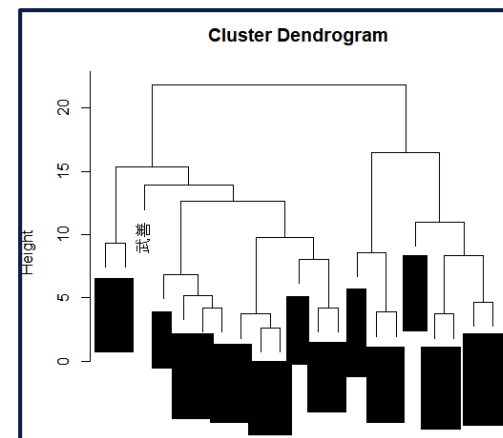
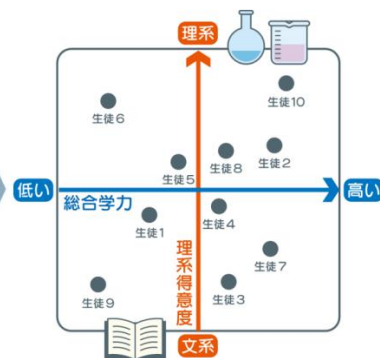
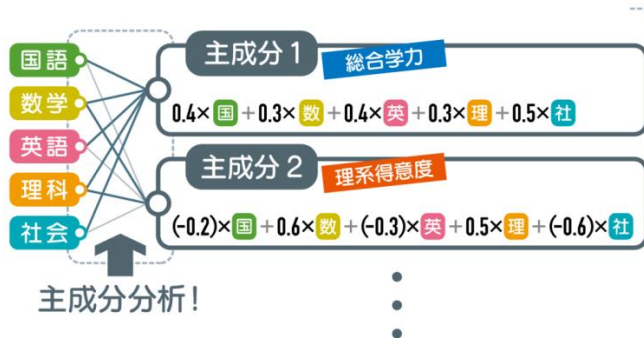
散布図行列

情報 II

そのまま見る!

主成分分析で、次元を圧縮しよう

クラスタリングで、グループ化しよう



etc...

情報Ⅰ／情報Ⅱ解説動画 プロジェクト

文科省「情報Ⅰ」「情報Ⅱ」解説動画 <https://www.nttls-edu.jp/joho/>

- ・ データサイエンス領域で情報Ⅰ、情報Ⅱともに関わった
- ・ 「これぞ情報Ⅰ」「これぞ情報Ⅱ」と言える内容にはなったと思われる

情報Ⅰ



情報Ⅱ



以降で、

情報Ⅰと情報Ⅱにおけるデータサイエンスの実践を紹介

目標

情報Ⅱ、案外出来そうだな

情報Ⅱ、やってみたくなる！

Agenda.

- 1 情報Ⅱのデータサイエンス(情報Ⅰと比較しながら)
- 2 情報Ⅱの授業演習**
- 3 情報Ⅰの逆向き設計
- 4 データサイエンスを扱う価値

まず、僕の基本的な「情報Ⅱ」授業の組み立て

- ・ 情報Ⅱは「探究」×「情報」の時間。
- ・ 「実践」→「理論」が基本の流れ(実習の後, 教科書を確認)。

「探究の種」(～2学期前半)

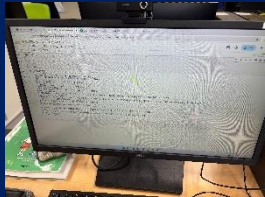
3D造形実習



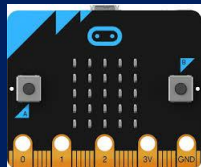
主成分分析実習



Python×AI実習



micro:bit 実習



1～3コマのパッケージ実習

etc...



卒業研究・卒業制作

2023 年度『情報の科学』卒業研究

■やること

※10/19(木)までに、Google Classroom ヘグループを登録する。

テーマ例一覧 (複数領域に跨る場合もある)

【推奨テーマ】

番号	テーマ	備考
A1	プログラミング等による開発(ハードウェア)	micro:bit 等を使用 (その他ロボット教材等あり)
A2	プログラミング等による開発(ソフトウェア)	ドリトル、Processing、表計算ソフト、Web ページなど。
B1	シミュレーション要素を含んだ調査研究	家具配置、ガチャシミュレーションの発展など
B2	データサイエンス要素を含んだ調査研究	データ分析 (アンケート調査、代表値計算、グラフ化) 等を行うこと
B3	情報技術・情報社会に関する調査研究	(例) 化粧の工夫は、同一人物判定の結果に寄与するか?

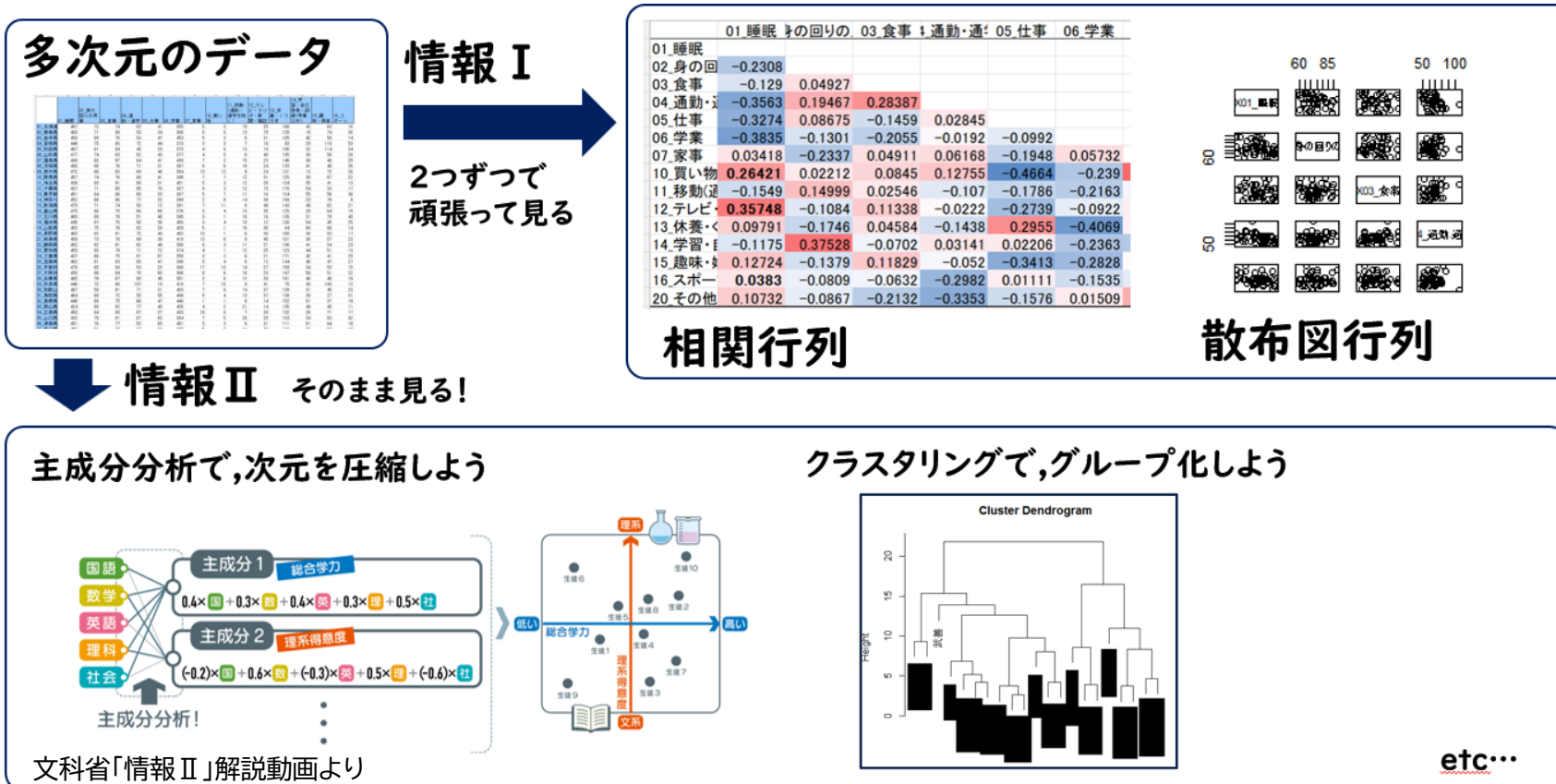


1～3人の
グループ実習



情報Ⅱ「データサイエンス領域」の指導の流れ

- ・ただし、統計分野だけは「理論」→「実践」。
- ・かつ体系的な指導にしている。
- ・ゴールは、卒業研究・卒業制作でのデータ分析のイメージが広がるように。



本日は、データサイエンス一連の授業の体験

2h

人工知能入門 モデルを用いた画像認識

人工知能(AI)とデータの関係性って何？

1h

クラスタリングで似ている友達を探せ

仕組みの概念的な理解を試してみる

3h

主成分分析で「好きなもののマップ」を作ろう

実際にデータサイエンスを使ってみる！

実は文科省動画をそのまま活用

2h

人工知能入門 モデルを用いた画像認識
人工知能(AI)とデータの関係性って何？

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



1h

クラスタリングで似ている友達を探せ
仕組みの概念的な理解を試みる

https://youtu.be/g2qkgKglA8c?si=uXPvN2AXS-_I_WwW



3h

主成分分析で「好きなもののマップ」を作ろう
実際にデータサイエンスを使ってみる！

https://youtu.be/rKwjOnzXw8U?si=sNuPzV_Rmo4J4Pv1



実習1

2h

人工知能入門 モデルを用いた画像認識
人工知能(AI)とデータの関係性って何？

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



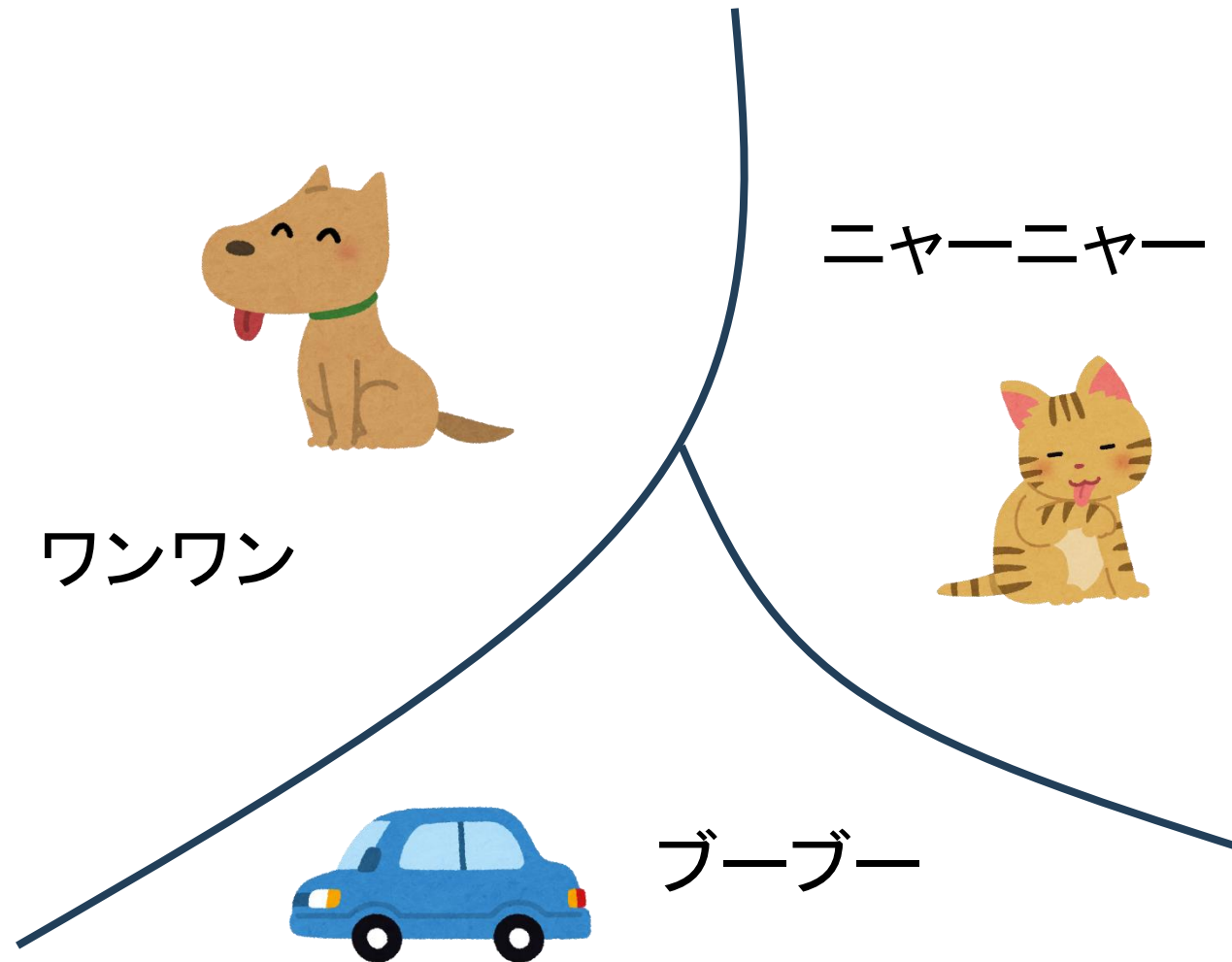
授業の流れ

- 今日から人工知能について勉強します！
 - 人工知能に関する話をいろいろする。
 - まずは一緒に動画を見てもらいます(～1:37くらいまで)
-

- 実は人工知能(AI)分野は, データサイエンスと関わりが深い。
- 人工知能はどんどん学習して賢くなる存在である(今の「顔」もそう)。
- 学習するってどういうこと？

ワンワン／ブーブー／ニャーニャーモデル

「四つ足の何か」空間



東京書籍情報Ⅱ
教科書挿絵より

「分かる」ことは、
「分ける」ことである！

(ミニ実習)Teachable Machineの体験

演習タイム

- ・「教師あり学習」を体験する
- ・このままScratchにエスカレートするのもOK。

ビデオの例で言えば…

学習の結果, 獲得したものが「顔」

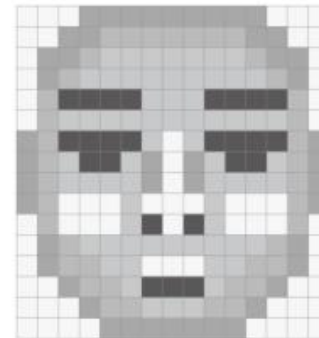
(学習時)

- ・「これが顔」という正解を与えて、学習を進めていく。

→「教師あり」学習



顔って、
こんなものだ!



学習フェーズ

学習画像



特徴量抽出

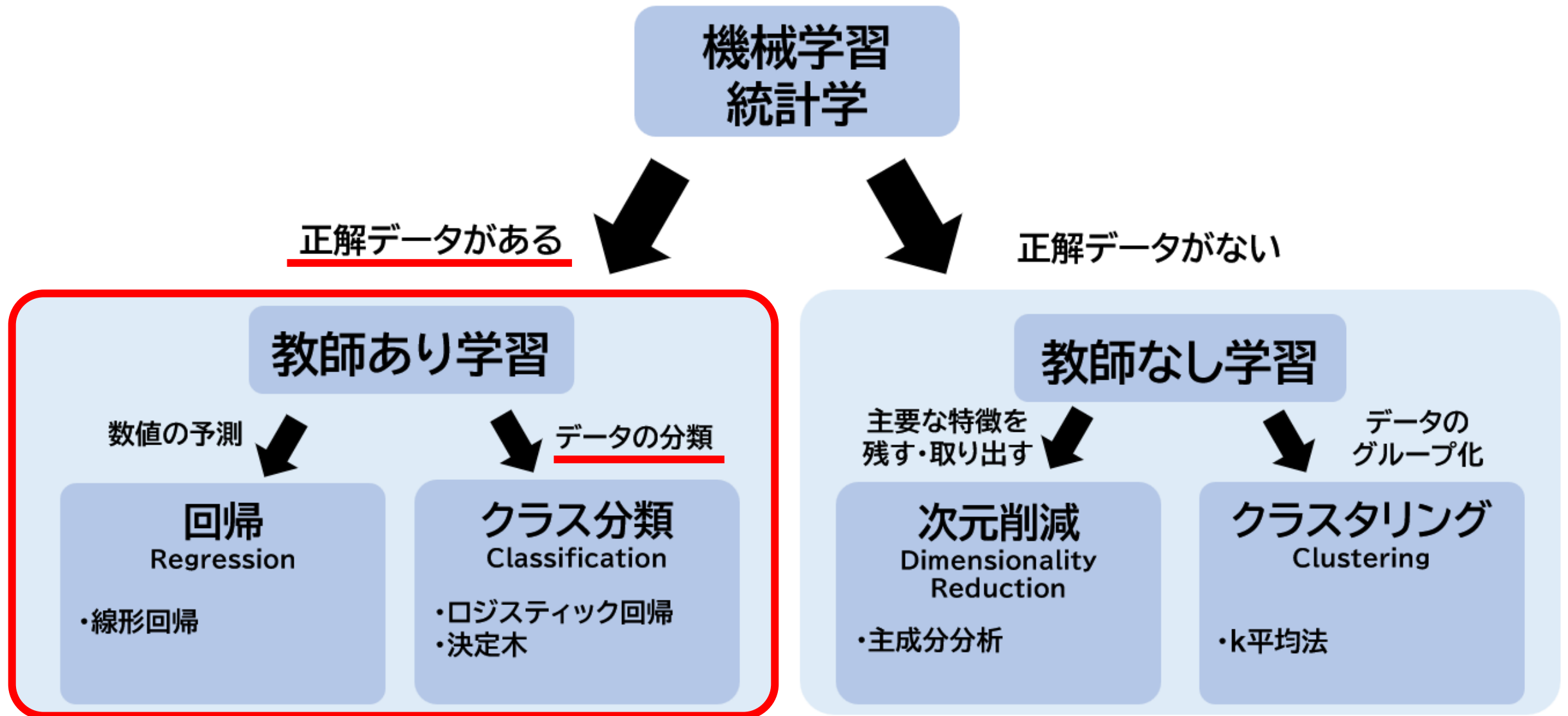


学習



学習結果データ

データサイエンスの全体像との対応



人工知能入門 モデルを用いた画像認識

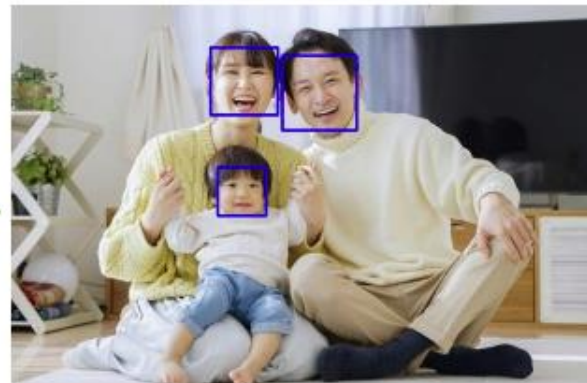
文科省「情報Ⅱ」解説動画より

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

この動画でやってみること

学習については別動画で！

学習済みのモデルを使って、 顔の検出・ぼかしに挑戦してみよう！



モデルを用いた画像認識

演習タイム

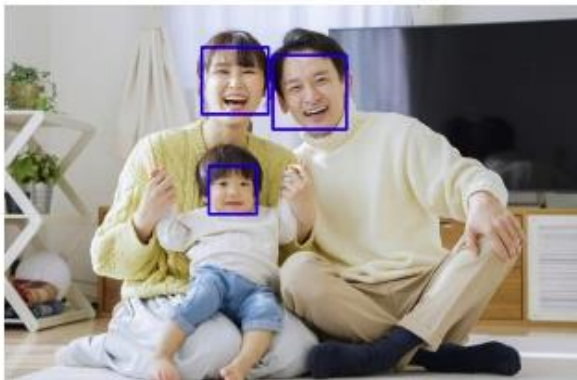
```
01 import cv2
02 from google.colab.patches import cv2_imshow
03 img = cv2.imread('syugou.jpg')
04 cv2_imshow(img)
05
06 face_cascade = cv2.CascadeClassifier('haarcascade_frontalface_alt.xml')
07 faces = face_cascade.detectMultiScale(img)
08
09 for(x,y,w,h) in faces:
10     face = img[y:y+h,x:x+w]
11     face = cv2.blur(face,(55,55))
12     img[y:y+h,x:x+w]=face
13 if len(faces)!=0:
14     cv2_imshow(img)
```

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



Google Colaboratoryを使用

文科省「情報Ⅱ」解説動画より



モデルを用いた画像認識 自動で顔にぼかしを入れよう！

生徒の実験・感想・考察

- aiで生成した画像の**キャラクターの目は認識された**のが面白かった。バーチャルアイドルなどのイラスト系はぜんぶダメでした。
- syamuのように**サングラスを掛けていると反応しません**でした→目のパーツを認識できないから？
- 月面にある顔に見えるものはAIには認識されなかった。
- 人ではできたかロボットでは顔として読み込むことができなかった
- 人の感情をaiが持てるのかも試してみたいです。

- モナリザなどの人の絵などは本物の絵も簡単に描いたイラストでも顔として認識された。しかし、ドラえもんや猫などは認識されなかった。
- ほかの人の画像認識の結果を聞いてもイラスト系の顔認識は判別しにくいことが分かった。
- **スカートの影を目として認識**していた。他にも髪の毛の分け目が目と認識された。

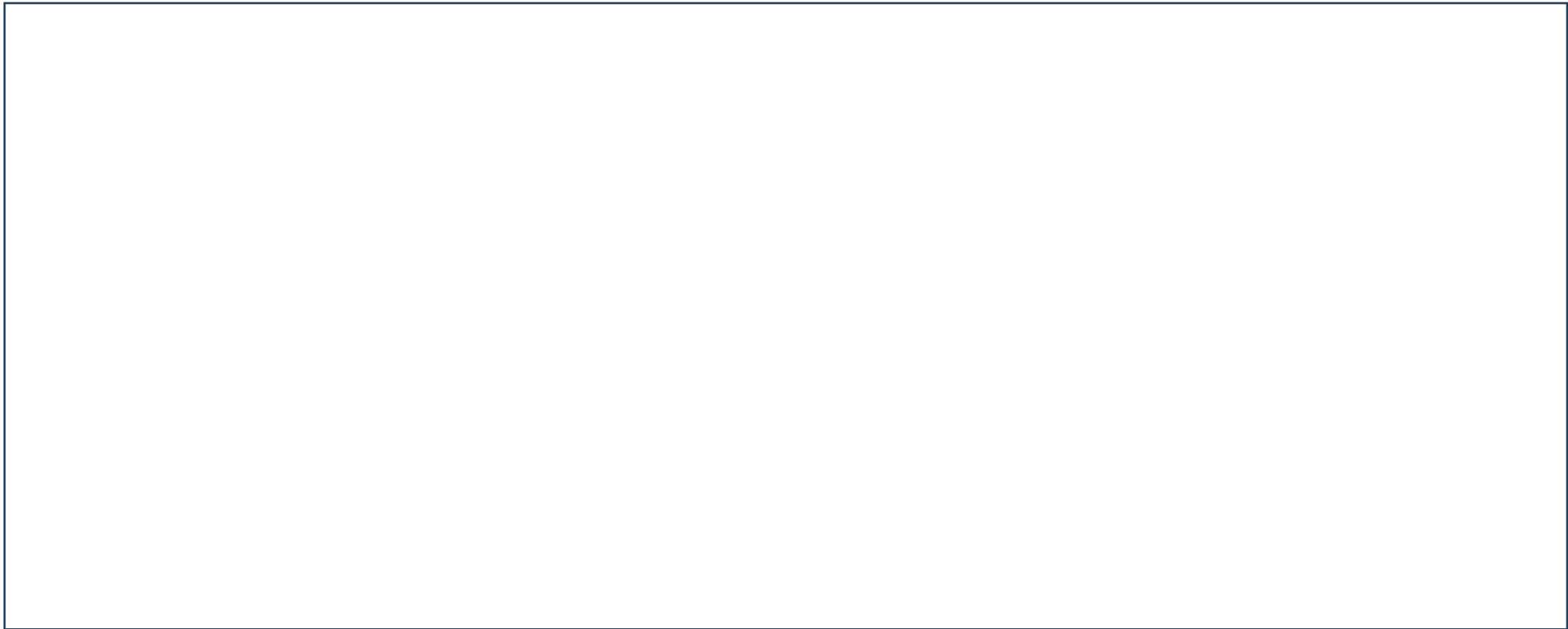
**AIの誤検出, 精度を上げる方法,
問題点に迫る気づきになる！**

(おまけの話) 人間の顔認知システム

日本心理学会 機関誌 心理学ワールド90号

<https://psych.or.jp/publication/world090/pw04/>

特集「乳幼児は顔を区別する」



新生児実験で使われた顔模式図形 (Morton & Johnson, 1991より)

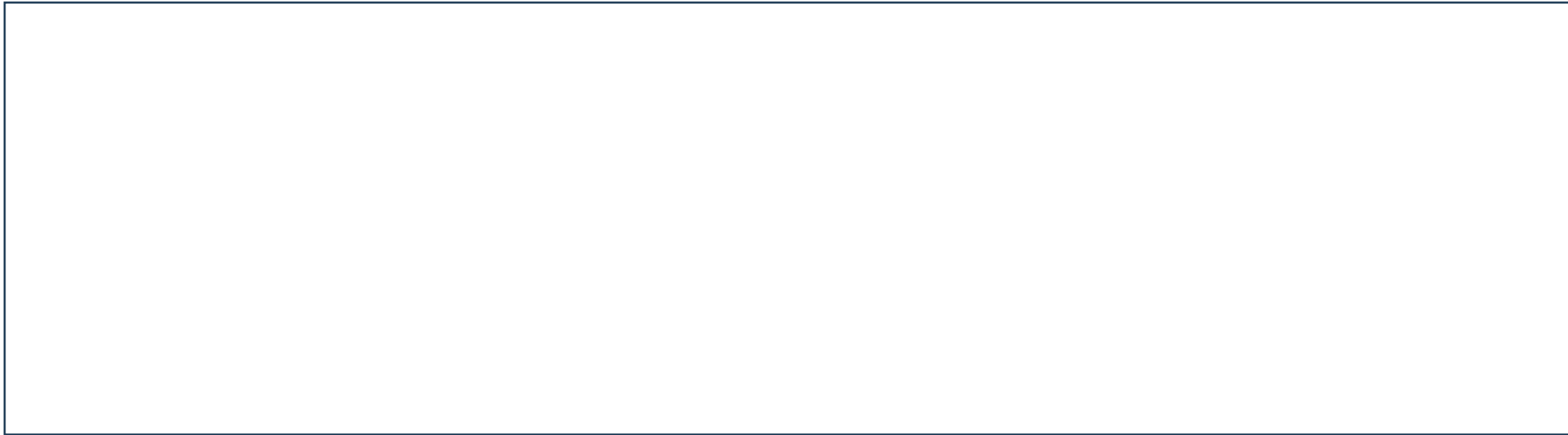
(おまけの話) 顔認知における文化差

日本心理学会 機関誌 心理学ワールド90号

<https://psych.or.jp/publication/world090/pw04/>

特集「乳幼児は顔を区別する」

目で表情をよむ日本人, 口で表情をよむ欧米人



日本の絵文字(上段)と欧米の絵文字(下段)

実習2

1h

クラスタリングで似ている友達を探せ
仕組みの概念的な理解を試みる

https://youtu.be/g2qkgKglA8c?si=uXPvN2AXS-_I_WwW



単元接続の流れ

- 人工知能のことを知るにも, データサイエンスが重要そうだ!
- 色んなモデルが大事なんだ!



- モデルってどうやって出来ているんだろう? 作るんだろう?
- 高校数学の範囲を逸脱せずに, イメージを掴む
+ 「教師あり学習」以外の統計手法を教える

「自分と近い性格の人を探そう!」実習

- ・まずはビデオを見てください(～3:00)

全体像に位置づけると

機械学習 統計学

正解データがある

正解データがない

教師あり学習

数値の予測

回帰
Regression

- ・線形回帰

データの分類

クラス分類
Classification

- ・ロジスティック回帰
- ・決定木

教師なし学習

主要な特徴を
残す・取り出す

次元削減
Dimensionality
Reduction

- ・主成分分析

データの
グループ化

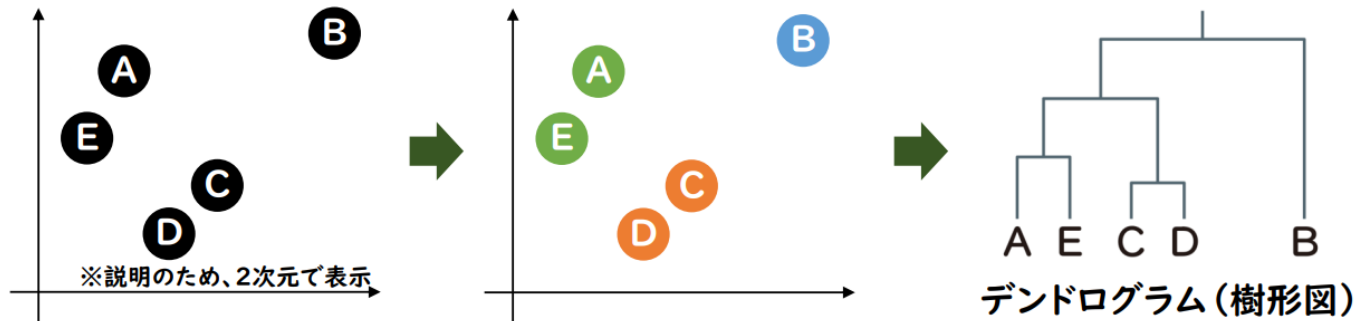
クラスタリング
Clustering

- ・k平均法

高校数学(中学数学)の範囲内でイメージ可能

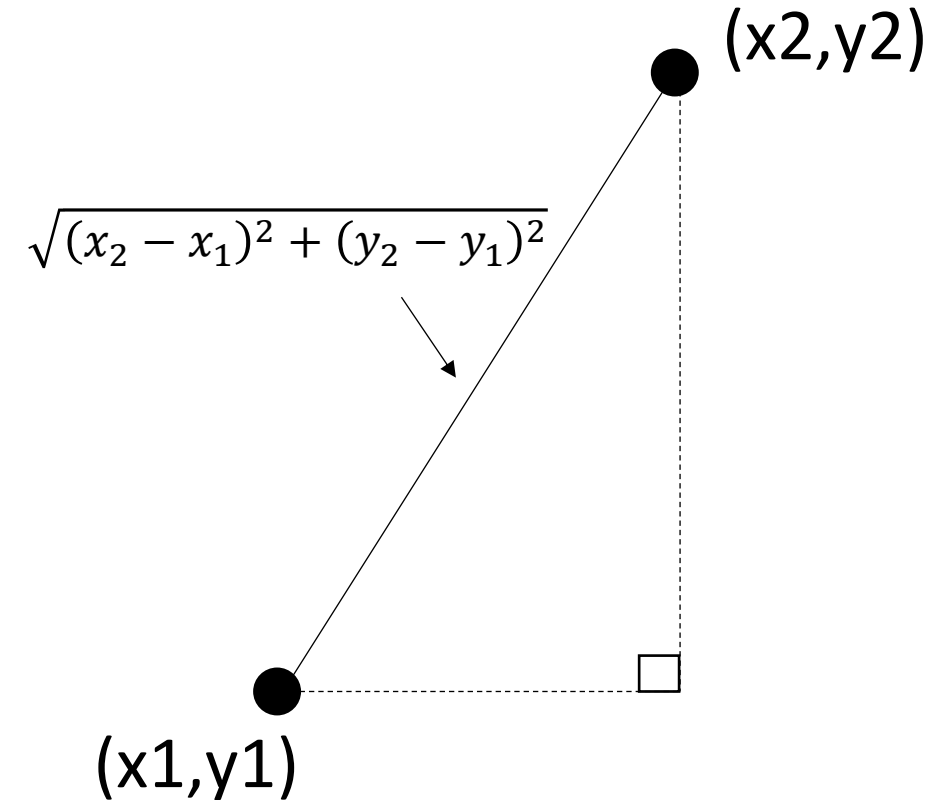
クラスタリング

データ間の距離を計算して、近いもの同士を結びつけて行く。



	洋楽	ドラマ	野球	JPOP	プログラライブ	バスケ	カラオケ	ペンギン	泳泳	登山	舞台鑑賞	パン	ダンス
生徒1	3	5	5	2	1	4	1	2	3	3	3	2	3
生徒2	3	3	3	2	3	1	5	1	2	5	3	5	5
生徒3	1	5	4	3	2	3	4	4	1	4	2	4	5
生徒4	4	3	5	2	1	3	5	4	1	2	2	2	4
生徒5	2	4	2	3	4	1	3	4	3	4	5	4	3
生徒6	3	5	1	2	5	2	5	5	2	1	3	5	2
生徒7	4	2	1	1	2	5	5	4	3	3	5	2	1
生徒8	2	3	3	3	1	3	1	3	4	3	4	5	5
生徒9	3	2	4	1	2	2	5	2	4	2	3	2	4
生徒10	4	5	4	2	1	2	4	3	3	4	4	4	3
生徒11	5	5	1	5	2	3	1	5	3	2	2	3	5
生徒12	5	1	1	5	5	4	1	5	4	2	1	4	3
生徒13	5	1	1	3	5	2	1	5	2	4	3	1	4
生徒14	4	2	4	1	3	1	2	3	4	2	3	5	3
生徒15	2	2	2	1	2	3	4	4	4	5	4	2	1

3次元以上の
データにも対応！



三平方の定理！

趣味を用いたクラスタリング

演習タイム

- (1) 共同編集スプレッドシートに1人1趣味入力
- (2) 1～5の5段階で「好き度」を記入
- (3) 記入されたシートを配信→Rでクラスタリング

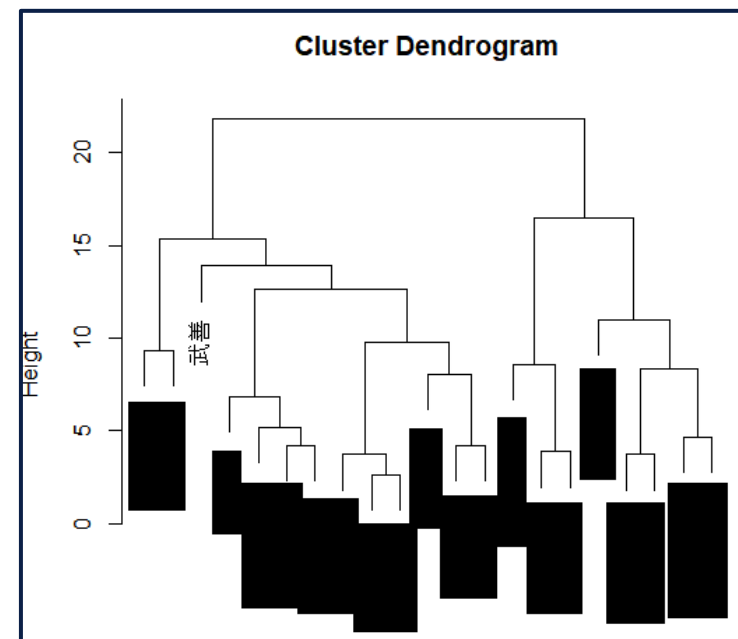
自分の名前の 下の欄に記入	ペン ギン	ゲー ム	サッ カー	睡眠	y o u t u b e	数学	水泳	野球	デ ィ ズ ニ ー	チ ョ コ	本	カー ビ ィ	車	バス ケ	ラー メン	散歩	運動
武善 紀之	5	5	2	5	3	4	4	2	3	4	5	4	2	2	4	5	3
	2	5	2	3	3	2	2	4	4	4	3	4	2	1	4	3	3
	4	5	5	4	5	2	4	2	3	4	2	4	1	2	5	3	4
	4	3	2	5	4	3	3	2	2	3	2	2	5	3	4	5	3
	4	4	5	5	5	2	3	2	4	5	2	5	3	3	5	5	4
	5	5	3	5	4	3	4	5	3	5	2	2	2	2	5	3	3
	2	4	3	5	5	2	5	3	5	3	3	3	2	3	4	3	3
	3	3	3	4	4	5	2	5	4	2	3	2	3	2	2	2	4
	2	4	1	5	5	3	4	2	5	5	4	2	2	2	5	5	2
	5	5	3	5	3	5	1	1	3	5	5	4	1	2	1	1	1
	2	5	2	5	5	5	1	1	3	4	5	1	2	1	5	2	1
	2	5	3	4	4	3	3	5	3	4	2	5	1	2	4	3	4
	1	5	2	5	5	4	5	5	3	2	1	3	5	4	5	4	5
	5	5	1	5	5	1	1	1	5	5	1	1	1	5	5	5	5
	3	4	2	5	5	3	3	4	3	5	2	2	2	2	5	5	4



<https://www.youtube.com/watch?v=g2qkgKglA8c>

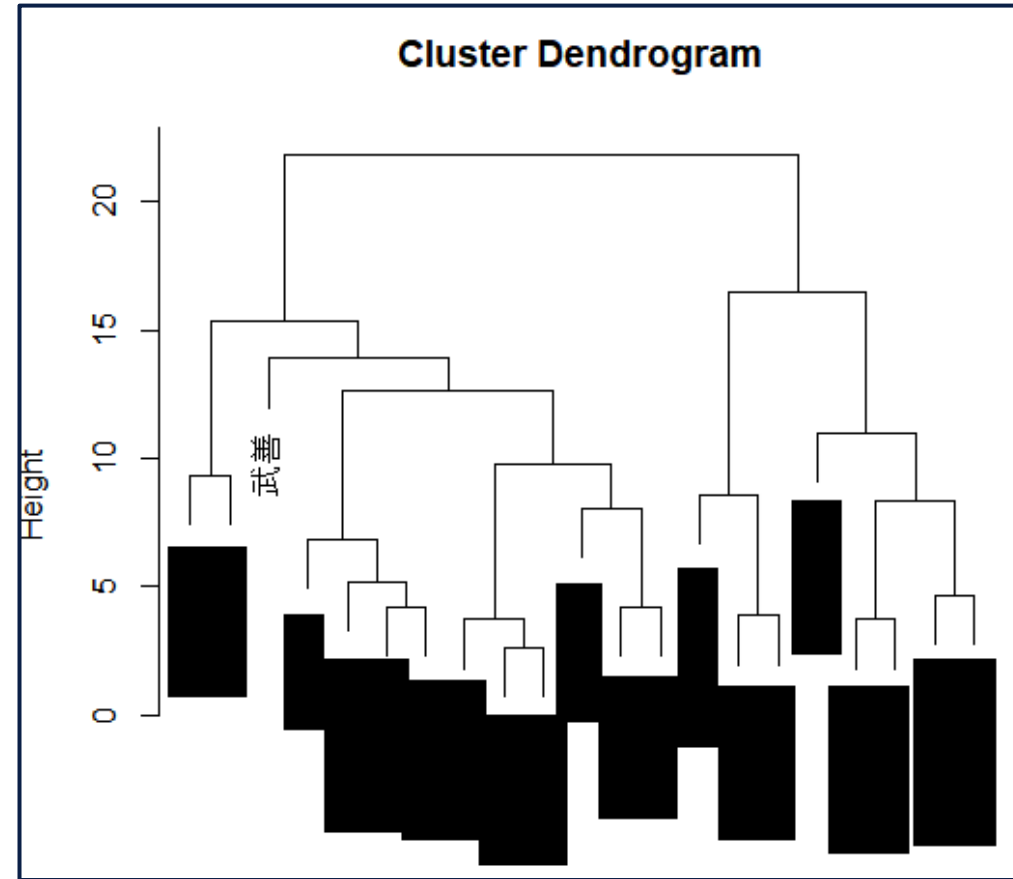


```
syumi <- read.csv("syumi.csv",h=T,row.names=1)
syumi.d <- dist(syumi)
syumi.c <- hclust(syumi.d)
plot(syumi.c)
```



クラスタリングの結果

<https://www.youtube.com/watch?v=g2qkgKglA8c>



生徒から本当に出てきた言葉

「でも、これって何に基づいているの…？」

この後の展開

- グループごとに集まってみる→解散
- グループごとの特徴を調べてみよう
- ビデオの続きを流す, その裏で僕は作業をする(1人TT)
- ビデオ 6:40~8:50くらい
- 今日はその作業の様子も同時並行でお見せします。

クラスタリングの感想

- ここまではっきり好きなもので別れるのは面白かった。数学がここまで情報に関係しているとは思わなかった。
- データ分析によっていろいろなものの特性を可視化することができてすごく面白いなと思った。多くの分野に応用することができそうだなと思った。
- マッチングアプリってこんな感じかーと思った。自分が数学で習ったものが実際に使われてとても関心した。
- あまり仲良くない人も同じグループにいたりして面白かった
- 自分のグループはほぼパソコン部の人だった。最後のその趣味が多いグループの表を見るのは面白かった
- 今話題になっているマッチングアプリやMBTIなどに応用されているのかなと考えました。

次の手法では

- 「なぜそのグループ」かを, クラスタリングでは自分で探る必要があった！
- 最初から, 「説明力」を持った手法もある。
- もう少し長い時間をかけて, 「自分の好き」を分類してみよう！

実習3

3h

主成分分析で「好きなもののマップ」を作ろう
実際にデータサイエンスを使ってみる！

https://youtu.be/rKwjOnzXw8U?si=sNuPzV_Rmo4J4Pv1



③自分の好きなものをマップ化→主成分分析

機械学習 統計学

正解データがある

正解データがない

教師あり学習

数値の予測

回帰

Regression

・線形回帰

データの分類

クラス分類

Classification

・ロジスティック回帰
・決定木

教師なし学習

主要な特徴を
残す・取り出す

次元削減

Dimensionality
Reduction

・主成分分析

データの
グループ化

クラスタリング

Clustering

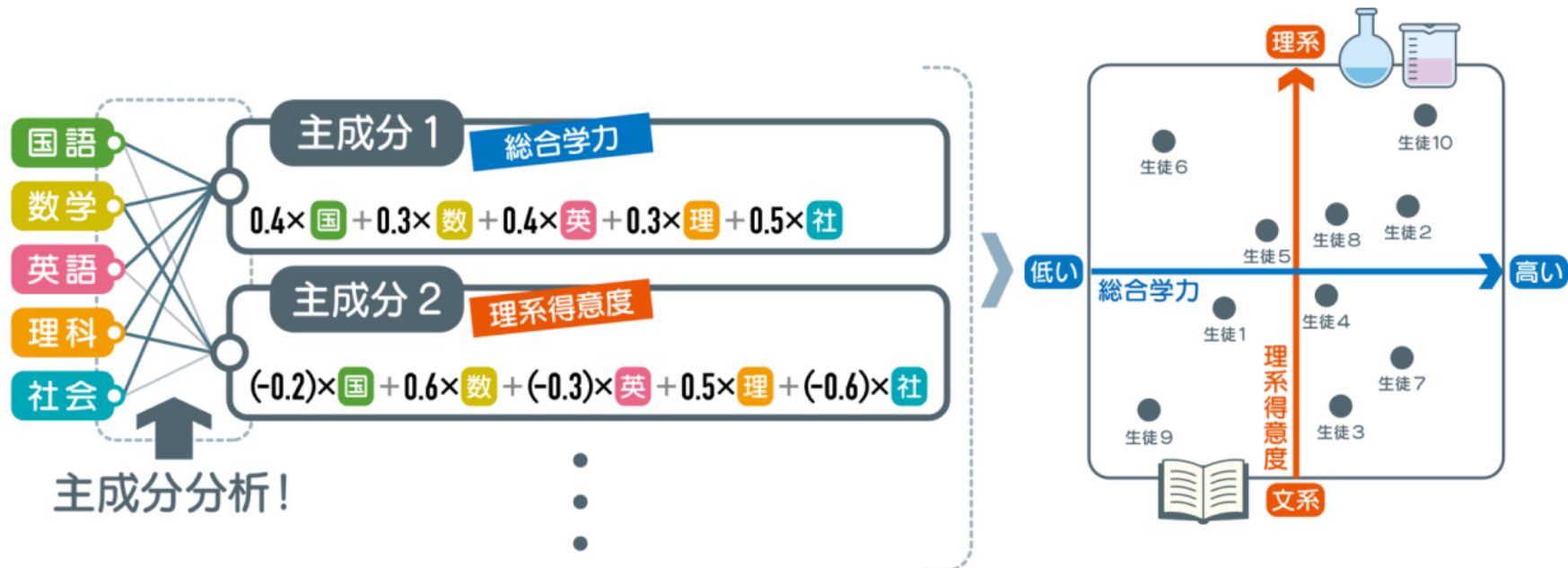
・k平均法

③自分の好きなものをマップ化→主成分分析

3番目にやる目的 いよいよ本格的に！ <https://www.youtube.com/watch?v=rKwjOnzXw8U>

主成分分析

多次元のデータを低次元の主成分に圧縮（要約）することで、データの構造を可視化できる。



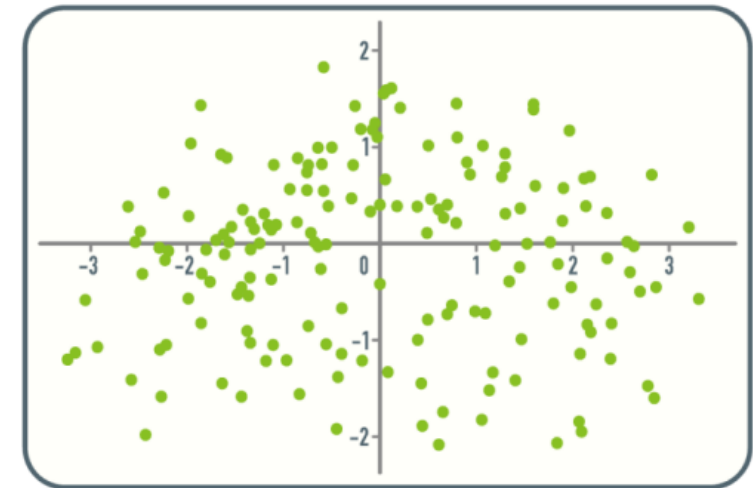
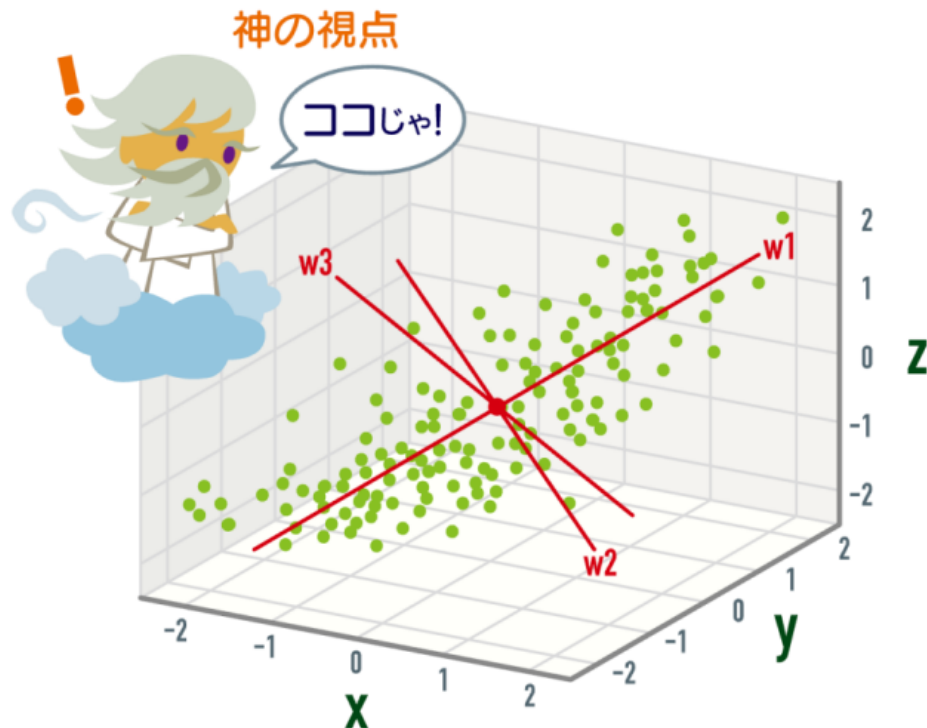
③自分の好きなものをマップ化→主成分分析

仕組み

<https://www.youtube.com/watch?v=rKwjOnzXw8U>

圧縮（要約）のイメージ

散らばりを最大化するように、空間上で軸を決める。
(一番散らばって見えるところで、見下ろすイメージ)



文科省「情報Ⅱ」解説動画より

ただし……

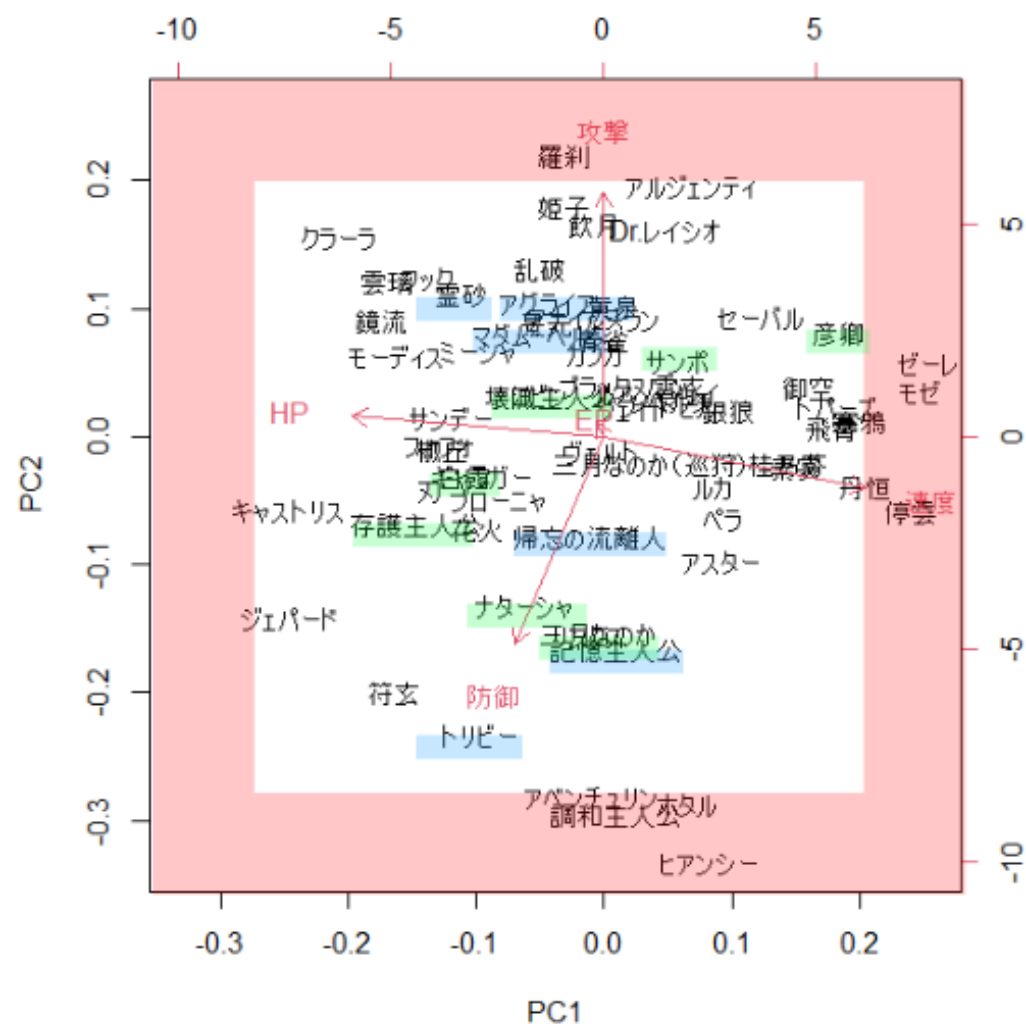
- 本日の講演では、扱いません。
- 理由：昨年度の研修で「主成分分析」は実施済み
- 授業での活用や流れについてだけ、ご説明します。

(演習タイム?)

- 



- ・ 2～3時間目: データ収集 & 分析



赤い範囲の中は比較的评价が高いとされるキャラ

评价が低いキャラ

评价が高いキャラ

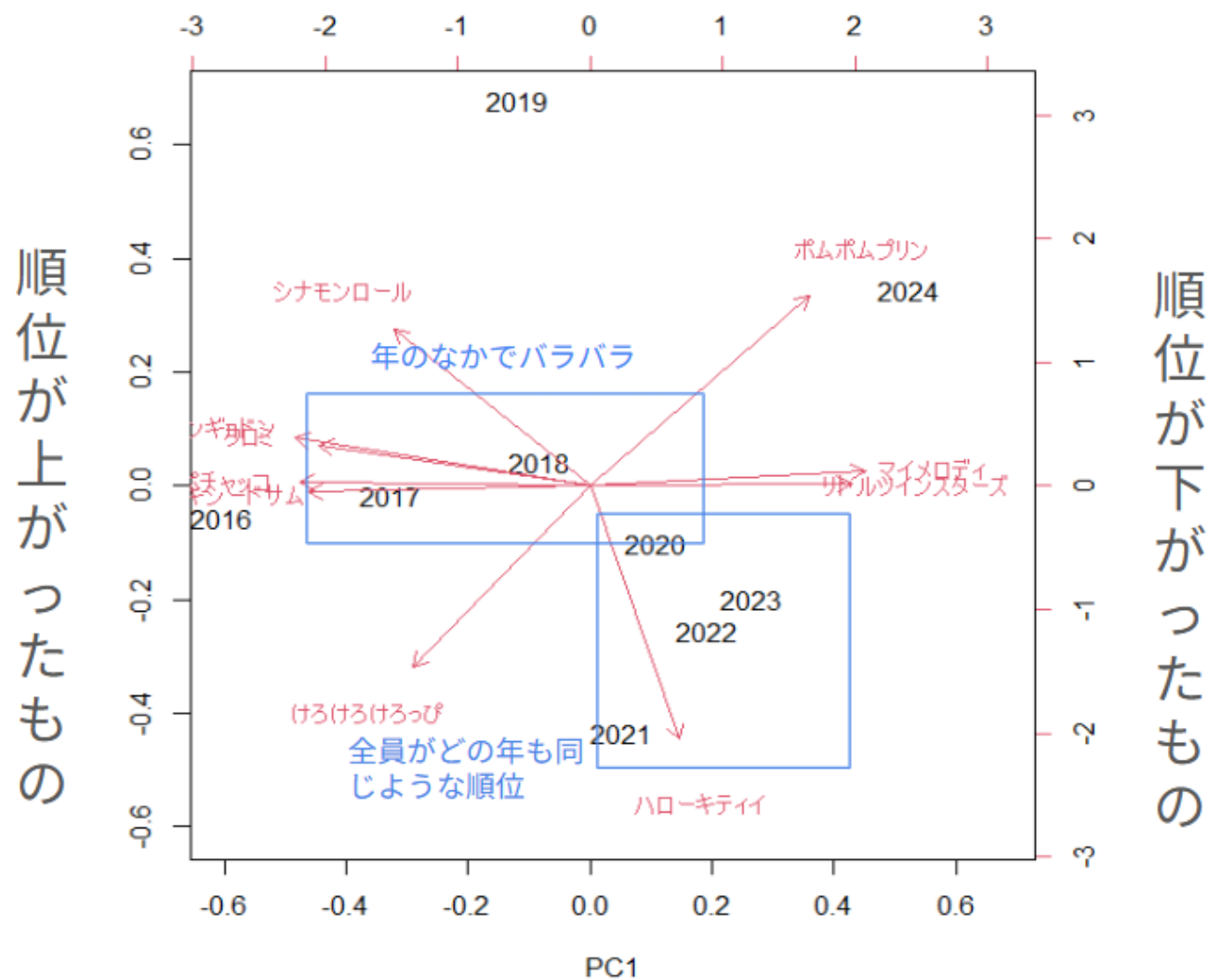
活用したデータ
崩坏スターレイルキャラ性能

使用した変数
攻撃 HP 防御 速度

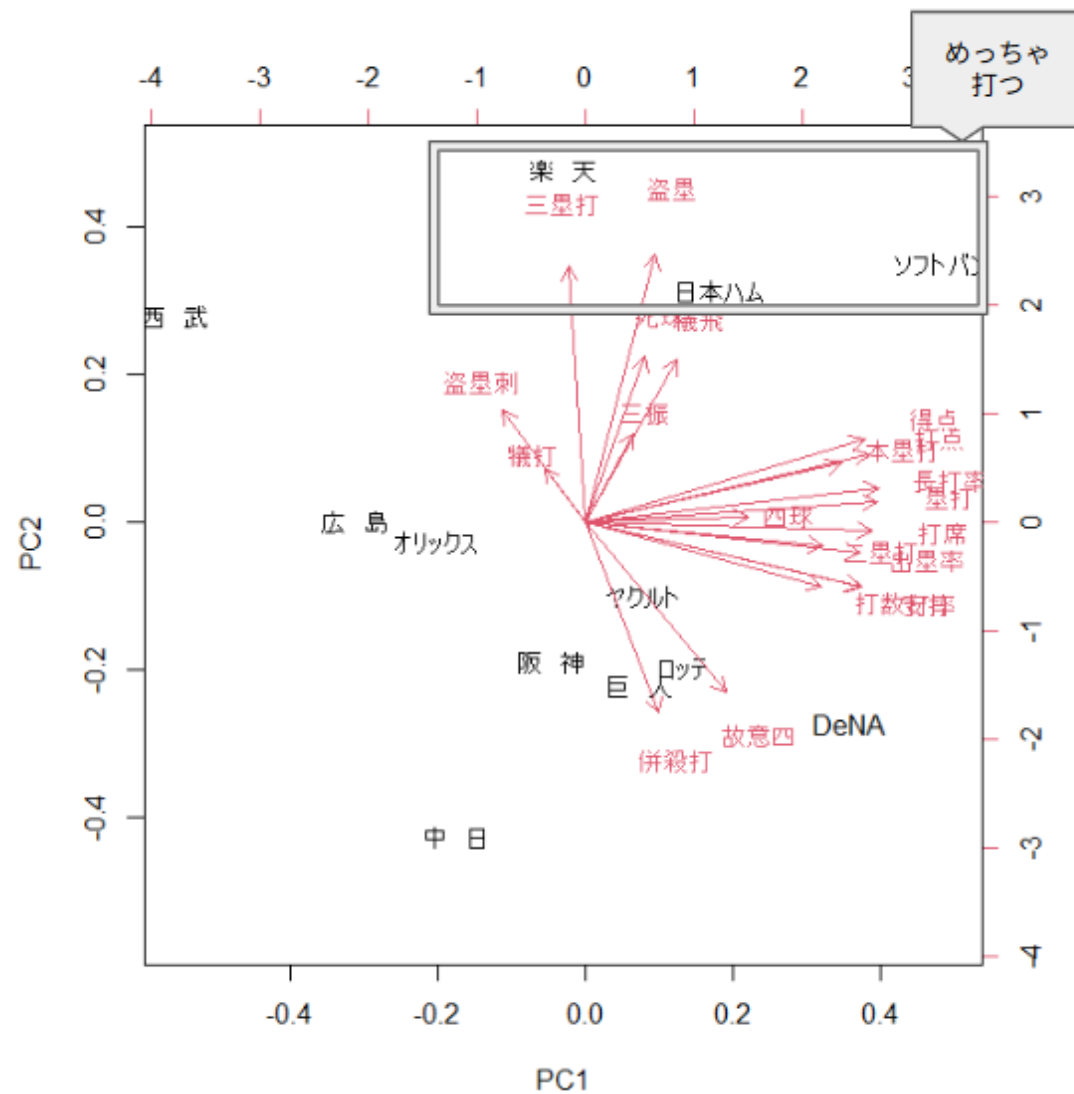
実習を通じて得られた知見

- 中心から外に行くにつれて、必ずプレイヤー評価が高い訳では無い
- 全体的にステータスがバラけていた

実際の生徒作品



実際の生徒作品



活用したデータ
NPBの2024年チーム成績

使用した変数
打率、打数、三振、盗塁など

実習を通じて得られた知見

-

以上で「データサイエンス」編終了

2h

人工知能入門 モデルを用いた画像認識
人工知能(AI)とデータの関係性って何？

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



1h

クラスタリングで似ている友達を探せ
仕組みの概念的な理解を試してみる

https://youtu.be/g2qkgKglA8c?si=uXPvN2AXS-_I_WwW



3h

主成分分析で「好きなもののマップ」を作ろう
実際にデータサイエンスを使ってみる！

https://youtu.be/rKwjOnzXw8U?si=sNuPzV_Rmo4J4Pv1



この実習を終えたあとの生徒達

この構図のイメージが、出来ているはず

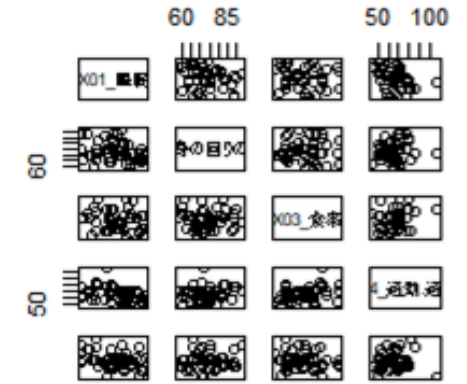
多次元のデータ

項目	1. 1週間(7日間)の生活時間								2. 1週間(7日間)の生活時間		3. 1週間(7日間)の生活時間		4. 1週間(7日間)の生活時間		5. 1週間(7日間)の生活時間		6. 1週間(7日間)の生活時間	
	起床	起床	起床	起床	起床	起床	起床	起床	起床	起床	起床	起床	起床	起床	起床	起床	起床	起床
01.起床	467	72	74	62	41	355	5	9	19	25	166	42	63	18	18	18	18	18
02.起床	459	71	80	53	34	386	5	9	19	25	166	42	63	18	18	18	18	18
03.起床	459	69	76	53	47	453	5	9	19	25	166	42	63	18	18	18	18	18
04.起床	449	75	60	72	44	373	5	9	19	25	166	42	63	18	18	18	18	18
05.起床	457	61	84	45	29	373	4	6	10	19	136	32	114	34	13	13	13	13
06.起床	477	74	83	52	49	377	4	5	8	6	40	135	38	58	28	28	28	28
07.起床	436	60	87	64	41	436	7	2	15	29	148	36	40	25	25	25	25	25
08.起床	436	60	87	64	41	436	7	2	15	29	148	36	40	25	25	25	25	25
09.起床	472	68	62	69	46	384	13	12	5	24	131	13	72	20	20	20	20	20
10.起床	457	74	78	68	41	386	7	2	12	31	125	36	67	22	22	22	22	22
11.起床	459	63	91	60	31	401	5	9	12	25	134	55	41	13	13	13	13	13
12.起床	437	71	85	65	70	387	6	3	12	15	170	54	33	17	17	17	17	17
13.起床	451	64	86	60	33	387	5	7	11	16	124	55	56	36	36	36	36	36
14.起床	452	69	86	77	32	386	5	8	14	39	108	23	78	9	9	9	9	9
15.起床	475	71	74	56	12	391	7	11	9	40	143	49	62	21	21	21	21	21
16.起床	475	66	75	66	68	376	3	4	10	35	125	29	64	15	15	15	15	15
17.起床	469	69	78	51	48	395	3	1	16	16	135	31	75	42	42	42	42	42
18.起床	445	67	75	56	38	400	3	4	10	12	129	54	45	25	25	25	25	25
19.起床	453	75	78	62	28	429	5	1	16	30	94	60	68	14	14	14	14	14
20.起床	463	62	81	72	43	402	10	1	9	26	150	39	53	17	17	17	17	17
21.起床	459	73	78	69	38	418	13	6	9	45	101	36	57	23	23	23	23	23
22.起床	452	62	81	62	40	390	9	1	11	21	136	41	54	23	23	23	23	23
23.起床	459	63	79	71	72	374	4	1	12	25	123	44	51	13	13	13	13	13
24.起床	431	61	87	67	59	3	6	6	21	171	156	42	41	23	23	23	23	23
25.起床	462	61	83	60	47	388	5	4	6	12	144	46	47	27	27	27	27	27
26.起床	476	65	83	53	23	360	17	10	14	27	159	34	53	15	15	15	15	15
27.起床	450	68	64	70	65	386	6	6	15	23	147	56	51	22	22	22	22	22
28.起床	465	79	87	69	45	351	3	6	12	33	161	40	48	19	19	19	19	19
29.起床	446	76	66	78	50	415	4	5	8	7	36	125	36	51	22	22	22	22
30.起床	467	56	91	71	31	402	7	6	14	37	139	31	45	23	23	23	23	23
31.起床	464	69	72	55	50	400	9	4	13	37	139	36	27	31	31	31	31	31
32.起床	449	68	75	66	47	440	1	1	9	14	152	51	37	19	19	19	19	19
33.起床	451	66	80	77	40	405	4	2	21	14	135	49	49	11	11	11	11	11
34.起床	459	64	80	67	71	433	18	6	7	20	152	39	71	17	17	17	17	17
35.起床	432	70	81	67	63	384	7	5	20	25	133	34	63	32	32	32	32	32
36.起床	451	76	77	62	65	401	3	3	3	8	31	111	61	64	18	18	18	18

情報Ⅰ

2つずつで
頑張ってる

	01 睡眠	02 身の回りの	03 食事	04 通勤・通学	05 仕事	06 学業
01 睡眠						
02 身の回りの	-0.2308					
03 食事	-0.129	0.04927				
04 通勤・通学	-0.3563	0.19467	0.28387			
05 仕事	-0.3274	0.08675	-0.1459	0.02845		
06 学業	-0.3835	-0.1301	-0.2055	-0.0192	-0.0992	
07 家事	0.03418	-0.2337	0.04911	0.06168	-0.1948	0.05732
10 買い物	0.26421	0.02212	0.0845	0.12755	-0.4664	-0.239
11 移動(通)	-0.1549	0.14999	0.02546	-0.107	-0.1786	-0.2163
12 テレビ	0.35748	-0.1084	0.11338	-0.0222	-0.2739	-0.0922
13 休養・趣味・娯楽	0.09791	-0.1746	0.04584	-0.1438	0.2955	-0.4069
14 学習・読書	-0.1175	0.37528	-0.0702	0.03141	0.02206	-0.2363
15 趣味・娯楽	0.12724	-0.1379	0.11829	-0.052	-0.3413	-0.2828
16 スポーツ	0.0383	-0.0809	-0.0632	-0.2982	0.01111	-0.1535
20 その他	0.10732	-0.0867	-0.2132	-0.3353	-0.1576	0.01509



相関行列

散布図行列

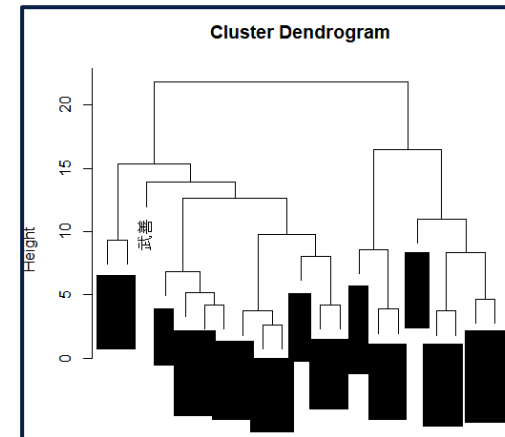
情報Ⅱ そのまま見る!

主成分分析で、次元を圧縮しよう



文科省「情報Ⅱ」解説動画より

クラスタリングで、グループ化しよう



etc...

学期末アンケートより

・ 情報Ⅱのデータサイエンス・統計領域を学んだ感想 （「数学Ⅰ」や「情報Ⅰ」との違いを中心に）

（全般的）

- ・ たくさんの項目をひとつにまとめて傾向や偏りが一目見てわかるようになるのはすごいと思った。
- ・ 主成分分析が人間が認識できないほど大きな次元をもとに分類できるのがすごいと思った。
- ・ 顔認識は写真アプリなどにも応用されているのかなと思い、普段の生活でも使われているのかもしれないと気づいた。
- ・ 情報Ⅰは覚えることや暗記がある程度多くてあまり得意ではありませんでしたが、情報Ⅱは実践的なものが多かったので楽しかったです。

（数学との比較）

- ・ 自分が数学で学んだ分散を使ってこのような主成分分析ができることを知り、とても驚きました。
- ・ 数学においてデータは二次元でまとめるものだと思っていたけど、主成分分析はいくつも軸があるやつをまとめられてこんな方法もあると知った
- ・ たくさんのデータをまとめることは数学などでは難しいので、それらをいっきにまとめられるのはとても便利だと思った。
- ・ 正確な値ではなく、傾向を見るのが新鮮だった。

見通しが立つように

機械学習 統計学

正解データがある

正解データがない

教師あり学習

数値の予測

回帰

Regression

- ・線形回帰

データの分類

クラス分類

Classification

- ・ロジスティック回帰
- ・決定木

教師なし学習

主要な特徴を
残す・取り出す

次元削減

Dimensionality
Reduction

- ・主成分分析

データの
グループ化

クラスタリング

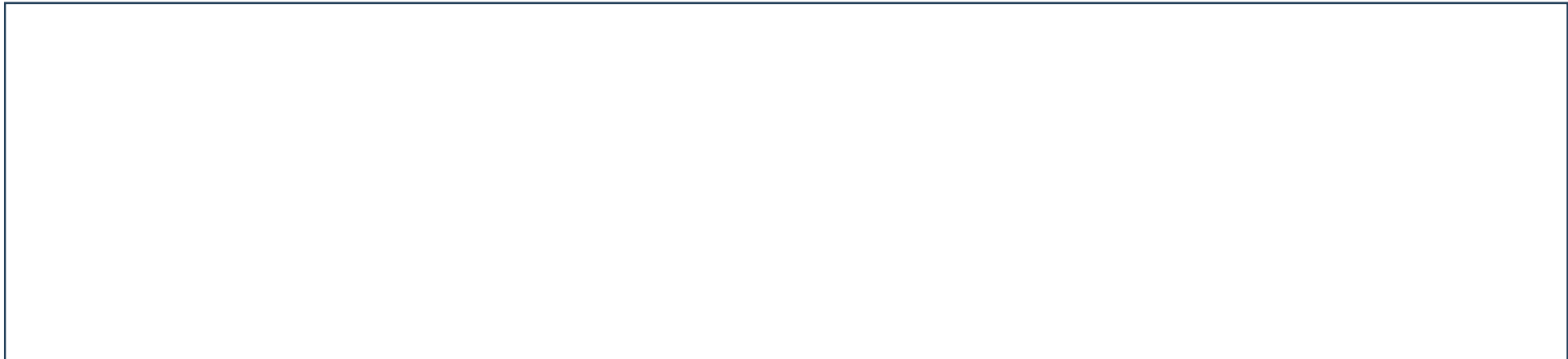
Clustering

- ・k平均法

第3章 情報とデータサイエンス

Theory12.機械学習と人工知能

- 計算機の処理能力の向上や機械学習、人工知能で用いられる手法の目覚ましい発展により、多様かつ多量の情報が適切に処理できるようになった。
- 機械学習には教師あり学習と教師なし学習がある。



今回は自分が出演した回を多く取り上げましたが

情報Ⅱ 解説動画

<https://www.nttls-edu.jp/joho/>

- 重回帰分析を用いた予測「睡眠時間を他の行動時間から予測しよう」
- 主成分分析による次元縮約「データを圧縮して、関係を見よう!」
- 機械学習による分類「手書きの数字をコンピュータに認識させよう」
- クラスタリング「自分と近い性格の人は誰？」
- ニューラルネットワークによる分類「より複雑な画像をコンピュータに認識させよう」
- モデルを用いた画像認識「自動で顔にぼかしを入れよう!」

情報とデータサイエンス



情報とデータサイエンスについて



重回帰分析を用いた予測
睡眠時間を他の行動時間から予測しよう



主成分分析による次元縮約
データを圧縮して関係を見よう!



機械学習による分類
手書きの数字をコンピュータに認識させよう



クラスタリング
自分と近い性格の人は誰?



ニューラルネットワークによる分類
より複雑な画像をコンピュータに認識させよう



モデルを用いた画像認識
自動で顔にぼかしを入れよう!

但し、ここまでの内容で終わっては駄目

- ・情報Ⅱは「探究」×「情報」の時間。
- ・「実践」→「理論」が基本の流れ(実習の後, 教科書を確認)。

「探究の種」(～2学期前半)

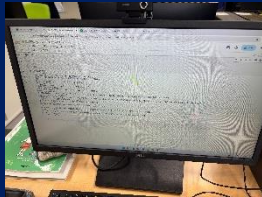
3D造形実習



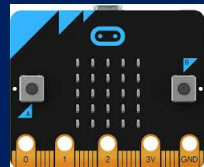
主成分分析実習



Python×AI実習

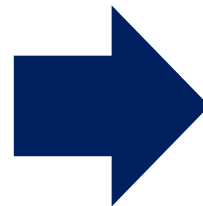


micro:bit 実習



1～3コマのパッケージ実習

etc...



卒業研究・卒業制作

2023 年度『情報の科学』卒業研究

■やること

※10/19(木)までに、Google Classroom ヘグループを登録する。

テーマ例一覧 (複数領域に跨る場合もある)

【推奨テーマ】

番号	テーマ	備考
A1	プログラミング等による開発(ハードウェア)	micro:bit 等を使用 (その他ロボット教材等あり)
A2	プログラミング等による開発(ソフトウェア)	ドリトル、Processing、表計算ソフト、Web ページなど。
B1	シミュレーション要素を含んだ調査研究	家具配置、ガチャシミュレーションの発展など
B2	データサイエンス要素を含んだ調査研究	データ分析 (アンケート調査、代表値計算、グラフ化) 等を行うこと
B3	情報技術・情報社会に関する調査研究	(例) 化粧の工夫は、同一人物判定の結果に寄与するか?



1～3人の
グループ実習



Agenda.

- 1 情報Ⅱのデータサイエンス(情報Ⅰと比較しながら)
- 2 情報Ⅱの授業演習
- 3 情報Ⅰの逆向き設計**
- 4 データサイエンスを扱う価値

この情報Ⅱに向かって、
情報Ⅰを設計する
(+入試)

情報Ⅰのデータサイエンス①

- 昔は色々やっていたが最近はともかく簡素に、最小限に。

仮説検定を高校生に教える

-情報科と数学科では何が違う？-



“怪しげな情報”の「作成・法則・検証」で
養うメディアリテラシー

「噂」と「広告」を題材に

- 「情報Ⅰ全体のボリュームUP」(社会と情報＋情報の科学)
- 「共通テスト対策への意識」(表計算の経験は必要)
- 「総合的な探究の時間」の実施(アンケート実習は総合で)
- 3変数以上は「情報Ⅱ」、2変数の処理を「情報Ⅰ」と分離

共通テスト「情報Ⅰ」試作問題

大学入試センター 試作問題「情報Ⅰ」より

第4問 次の文章を読み、後の問い（問1～5）に答えよ。（配点 25）

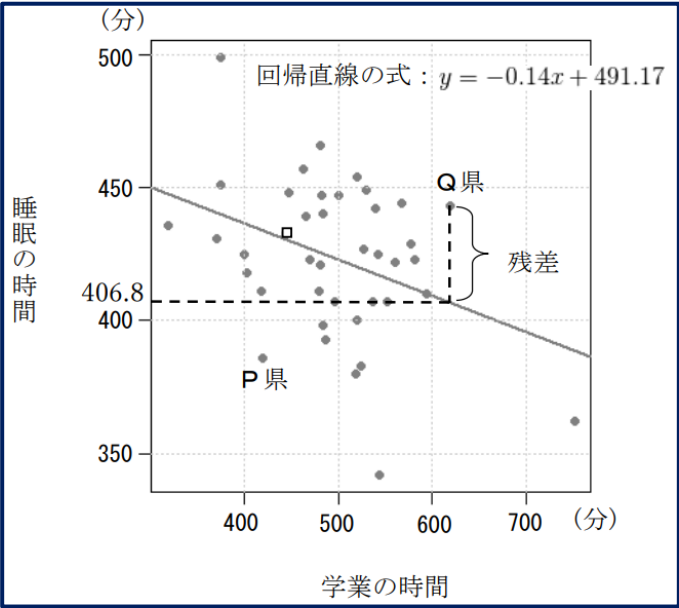
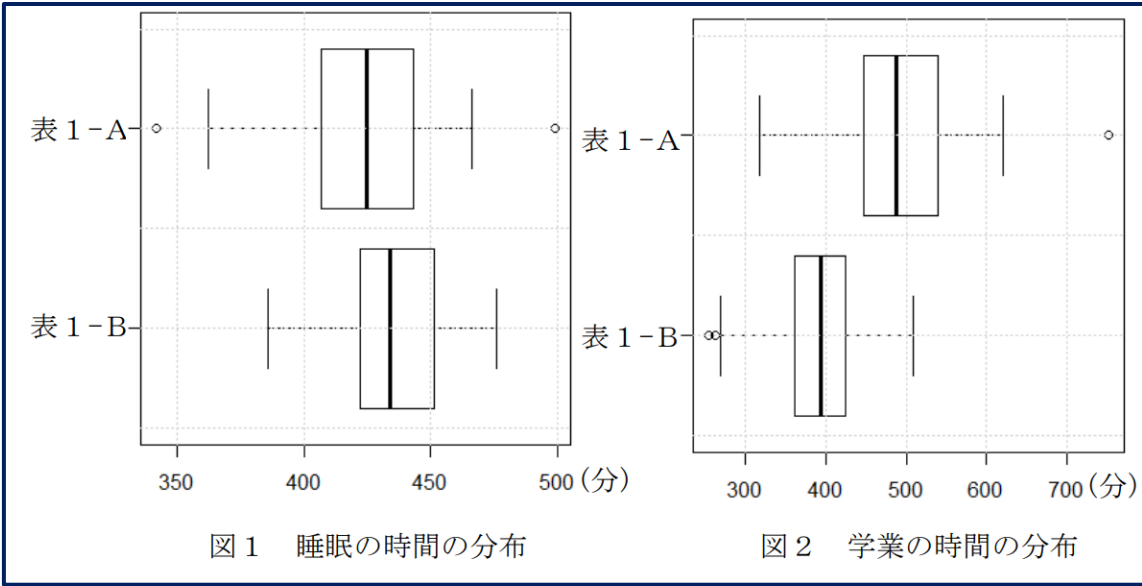
次の表1は、国が実施した生活時間の実態に関する統計調査をもとに、15歳以上19歳以下の若年層について、都道府県別に平日1日の中で各生活行動に費やした時間（分）の平均値を、スマートフォン・パソコンなどの使用時間をもとにグループに分けてまとめたものの一部である。ここでは、1日のスマートフォン・パソコンなどの使用時間が1時間未満の人を表1-A、3時間以上6時間未満の人を表1-Bとしている。

表1-A：スマートフォン・パソコンなどの使用時間が

1時間未満の人の生活行動時間に関する都道府県別平均値

都道府県	睡眠 (分)	身の回りの 用事 (分)	食事 (分)	通学 (分)	学業 (分)	趣味・娯楽 (分)
北海道	439	74	79	60	465	8
青森県	411	74	73	98	480	13
茨城県	407	61	80	79	552	11
栃木県	433	76	113	50	445	57

- 多次元の表, 箱ひげ図, 散布図は必須
- 短時間で

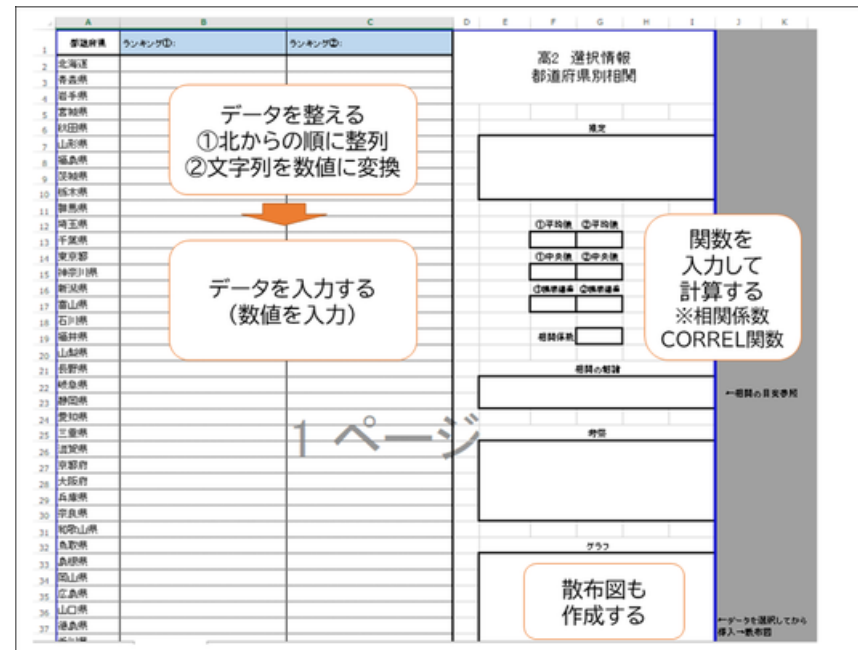
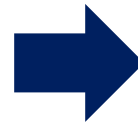


「これだ！」と思えた実践

第16回全国高等学校情報教育研究会全国大会(東京大会)

「とどらん(都道府県ランキング)で教えるデータ活用の授業」

アサンプション国際中学校高等学校 岡本弘之先生

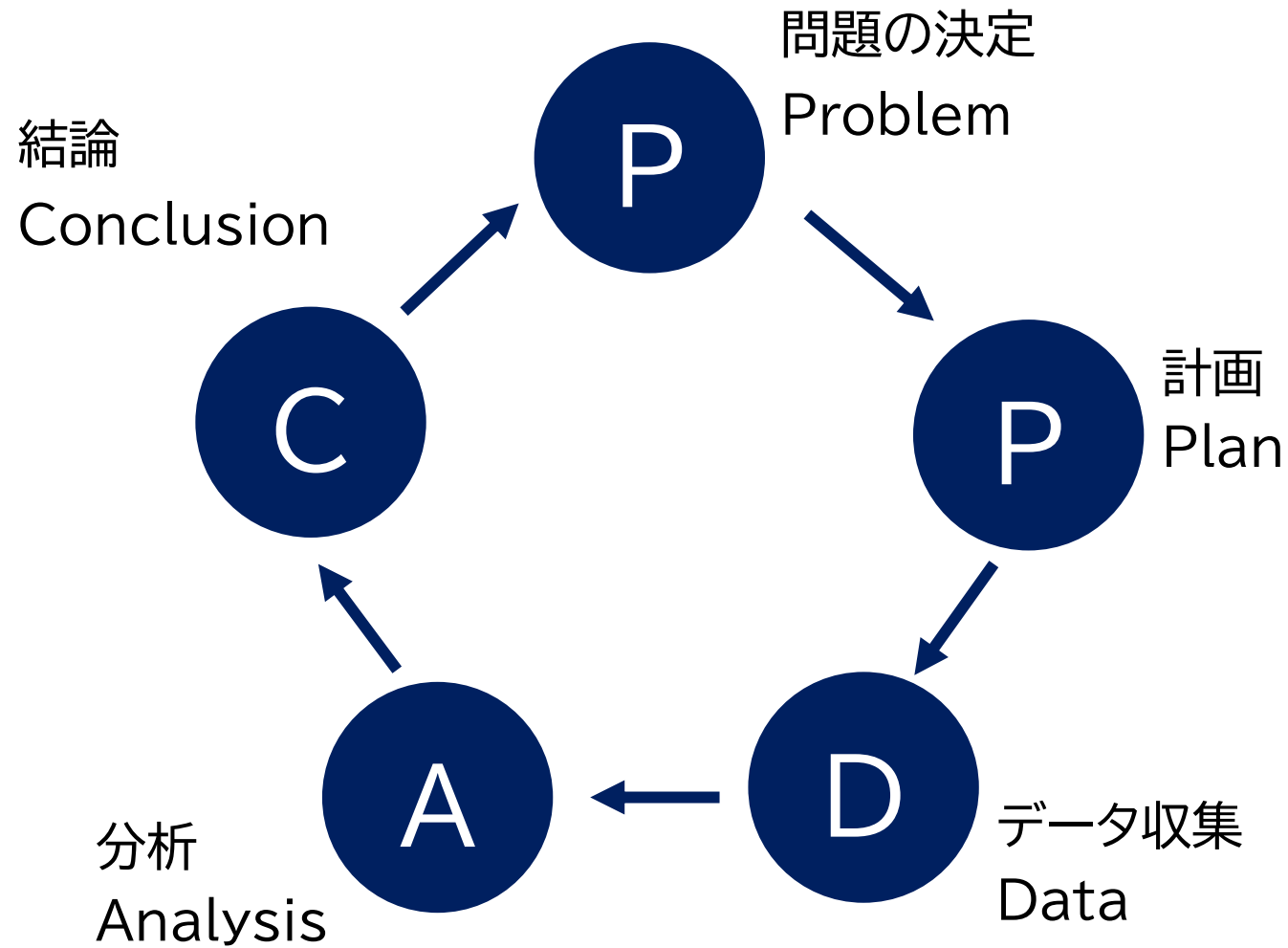


<https://todo-ran.com/>

(画像)キミの未来発見事例205

<https://www.wakuwaku-catch.net/jirei22205/>

PPDACサイクルも意識されている



3時間で実施(データサイエンス入門)

(1時間目)

・数学復習&操作習得

(2時間目)

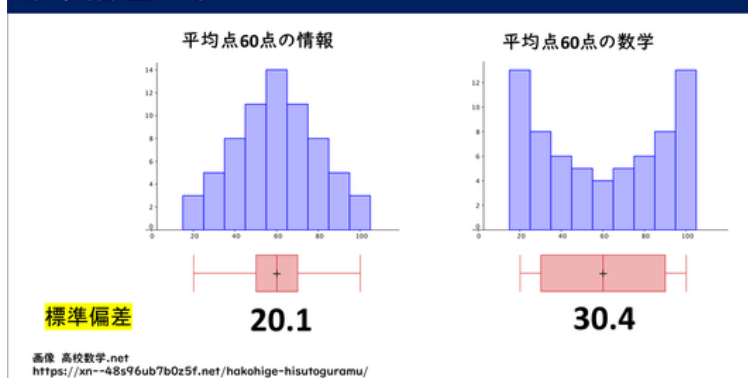
・各自の分析&発表

(3時間目)

・振り返り分析

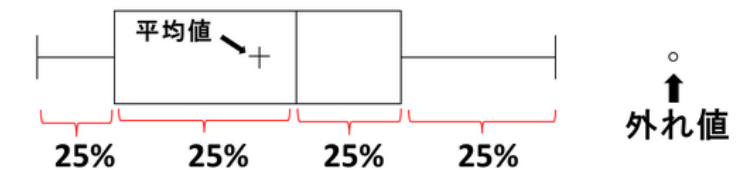
数学の復習(ヒストグラム・箱ひげ図・相関係数)

標準偏差を求めると...



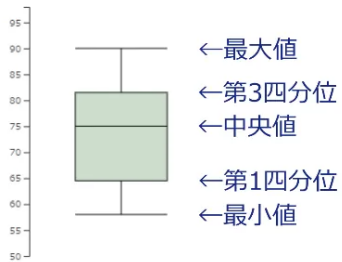
箱ひげ図を書いてみる!

箱ひげ図のポイントはこれだけ!



箱ヒゲ図作成サイト

統計データをあらわす時に、箱ヒゲ図を使うことがあります。箱ひげ図の書き方は、以下の図のように最大値、第3四分位、中央値、第1四分位、最小値を箱の形で示します。

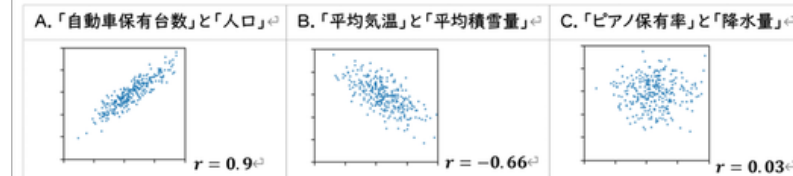


上図のように、箱があり、その両端にヒゲがあるように見えますね。

- ・箱の部分：全体の1/4から3/4まで
- ・ヒゲの部分：最小値から最大値まで

と、データの範囲が分かりやすく図示されます。

相関係数



-1.0 ~ -0.7	-0.7 ~ -0.4	-0.4 ~ -0.2	-0.2 ~ 0.2	0.2 ~ 0.4	0.4 ~ 0.7	0.7 ~ 1.0
強い負の相関	負の相関	弱い負の相関	相関なし	弱い正の相関	正の相関	強い正の相関

-1 ~ 1の値をとる。

画像: のんびり統計
https://note.com/e_doo/n/n7697b32f1a12

☆箱ひげ図も作成させる

<https://statistics.calculator.jp/basic/box-plot/>

3時間で実施(データサイエンス入門)

(1時間目)

- ・数学復習&操作習得

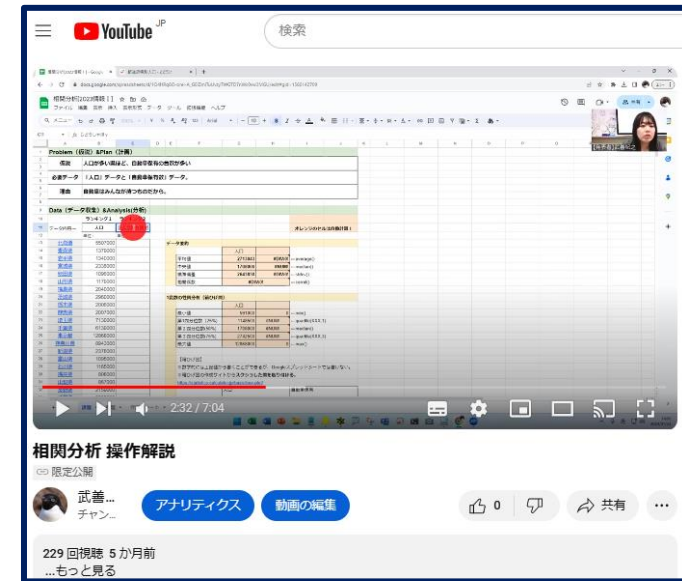
(2時間目)

- ・各自の分析&発表

(3時間目)

- ・振り返り分析

YouTubeで操作解説動画



共同編集スプレッドシートで知見共有

氏名	仮説	データ1	データ2	相関係数	次に検証したい仮説
武善 紀之	人口が多いほど、自動車を保有	人口	自動車保有台数	0.95	ガソリンスタンドの数は面積と人口、どちらに関係するか
	高齢者が多いほど、パチンコ店	人口	パチンコ店舗数	0.985	人口とパチンコ店舗数の関係
	都会になるにつれて家賃が高く	人口	家賃	-0.098	サッカーと野球の人口
	睡眠時間が短いほど年収が高い	平均年収	平均睡眠時間	-0.735	大企業数が多いと年収が高い
	降雪量が多いとスキーやスノー	競技人口	降雪量	-0.003	降雪量とスキーやスノーボードの競技人口競技人口の県別
	生活保護と人口の比例	生活保護世帯	人口	0.918	テニスと卓球の人口
	魅力度が高いほどふるさと納税	ふるさと納税額	魅力度	0.432	漁獲量とふるさと納税
	人口が多い都道府県のほうがセ	セブンの数	人口	0.967	セブンの数とファミマの数
	お菓子消費量が多いほど市街地	お菓子消費量	市街地比率	0.705	食べ物の消費量と市街地比率

3時間で実施(データサイエンス入門)

(1時間目)

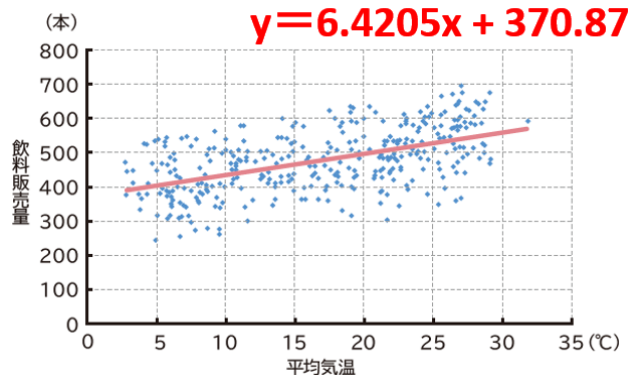
- ・数学復習&操作習得

(2時間目)

- ・各自の分析&発表

(3時間目)

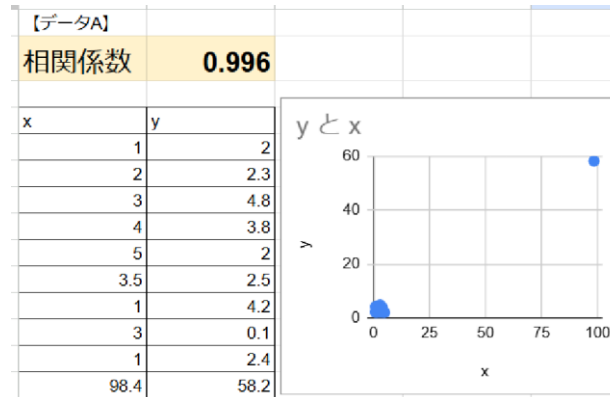
- ・振り返り分析



検討1.相関係数は本当に正しい？

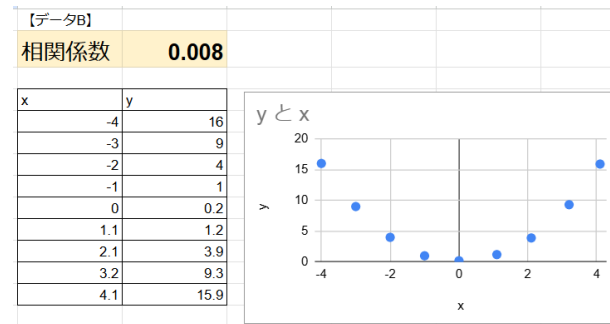
・相関係数 0.98

すごい関係性あるな！

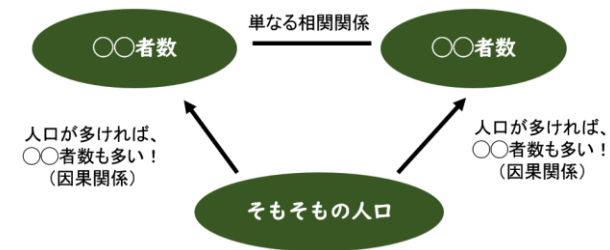


振り返ろう2

自分が調べたものは、相関関係か？
因果関係か？因果の方向性はどちらか？他に考えられる要因は何か？



この関係性に陥っている人が多いそう。



振り返ろう3 (分析手法の発展)

単回帰分析の
数式が意味することを書いてみよう。

一応この内容で「情報Ⅰ」共通テストは網羅

大学入試センター 試作問題「情報Ⅰ」より

第4問 次の文章を読み、後の問い（問1～5）に答えよ。（配点 25）

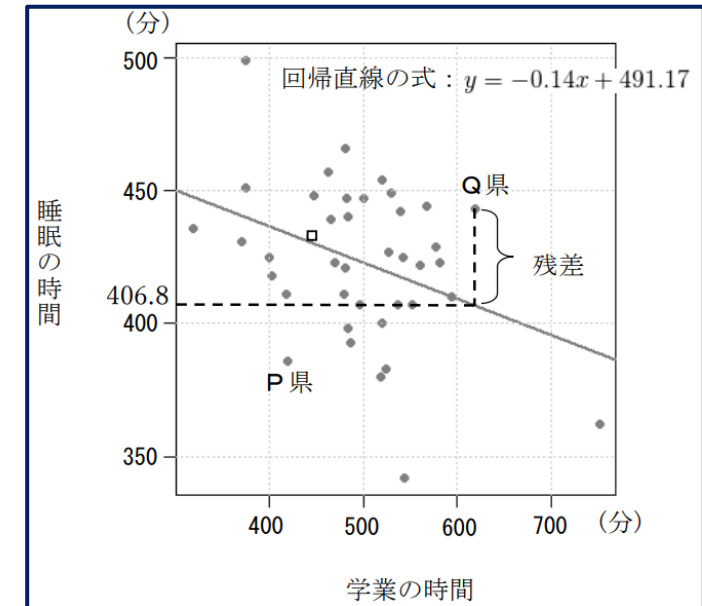
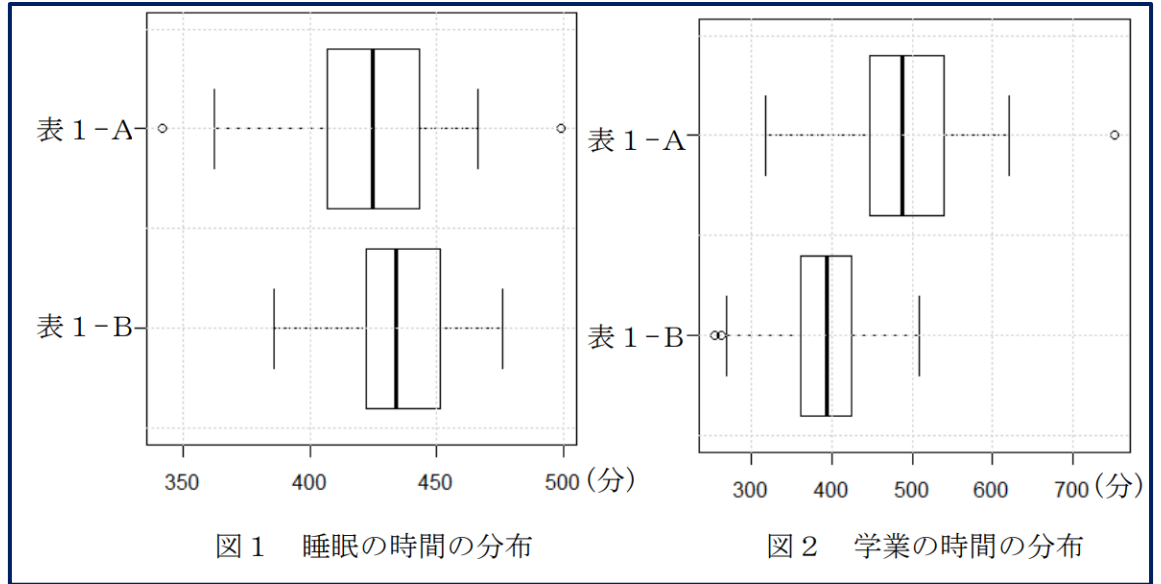
次の表1は、国が実施した生活時間の実態に関する統計調査をもとに、15歳以上19歳以下の若年層について、都道府県別に平日1日の中で各生活行動に費やした時間（分）の平均値を、スマートフォン・パソコンなどの使用時間をもとにグループに分けてまとめたものの一部である。ここでは、1日のスマートフォン・パソコンなどの使用時間が1時間未満の人を表1-A、3時間以上6時間未満の人を表1-Bとしている。

表1-A：スマートフォン・パソコンなどの使用時間が

1時間未満の人の生活行動時間に関する都道府県別平均値

都道府県	睡眠 (分)	身の回りの 用事 (分)	食事 (分)	通学 (分)	学業 (分)	趣味・娯楽 (分)
北海道	439	74	79	60	465	8
青森県	411	74	73	98	480	13
茨城県	407	61	80	79	552	11
栃木県	433	76	113	50	445	57

- 多次元の表, 箱ひげ図, 散布図は必須
- 短時間で



情報Ⅰ／情報Ⅱで培われる視点

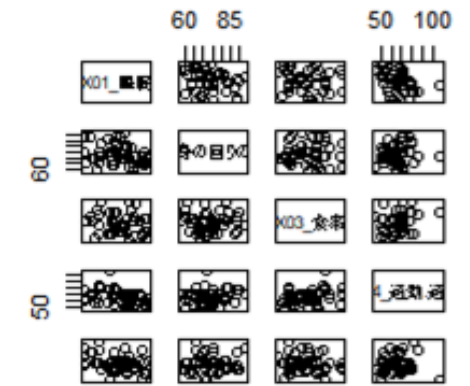
多次元のデータ

	01.睡眠	02.身の回りのこと	03.食事	04.通勤・通学	05.仕事	06.学業	07.家事	08.買いもの	09.移動・交通	10.趣味・娯楽	11.休養・くつろぎ	12.テレビ・動画・音楽	13.読書・新聞	14.学習・自己啓発	15.その他
01.睡眠	467	72	74	62	41	355	5	9	19	25	166	42	63	18	
02.身の回りのこと	469	71	80	53	34	386	5	9	15	79	129	19	14	30	
03.食事	459	69	76	53	47	453	5	9	9	31	105	42	59	14	
04.通勤・通学	449	65	72	44	170	3	5	1	15	50	39	119	53		
05.仕事	467	61	84	45	29	373	4	6	10	19	136	32	114	34	
06.学業	477	63	82	40	377	4	5	6	40	135	39	59	20		
07.家事	439	60	87	64	41	439	7	2	15	29	149	36	40	25	
08.買いもの	439	60	87	64	41	439	7	2	15	29	149	36	40	25	
09.移動・交通	472	66	82	69	46	384	13	12	5	24	131	13	13	20	
10.趣味・娯楽	457	74	78	69	41	398	7	7	12	31	125	36	67	22	
11.休養・くつろぎ	459	63	91	60	31	401	5	9	12	25	154	55	41	13	
12.テレビ・動画・音楽	437	71	85	65	70	387	6	9	12	15	170	54	33	17	
13.読書・新聞	451	64	86	69	33	387	5	7	11	16	124	55	56	36	
14.学習・自己啓発	452	69	86	77	32	389	5	8	14	39	108	23	79	9	
15.その他	475	71	74	56	12	391	7	11	9	40	143	49	62	21	
16.睡眠	475	66	75	66	68	376	3	4	10	35	125	29	64	15	
17.身の回りのこと	469	69	78	51	48	395	3	1	16	16	135	31	75	45	
18.食事	445	67	75	56	38	450	3	4	10	12	129	54	45	25	
19.通勤・通学	453	75	78	62	28	429	5	1	16	30	94	60	68	14	
20.仕事	463	62	81	72	43	402	10	1	9	19	150	39	53	17	
21.学業	459	73	79	69	39	418	13	6	9	45	101	36	57	23	
22.家事	452	62	81	62	49	390	9	9	11	13	136	41	54	23	
23.買いもの	459	63	79	71	72	374	4	1	12	25	123	44	51	13	
24.移動・交通	471	66	81	67	59	399	5	6	11	17	111	42	41	23	
25.趣味・娯楽	462	61	83	60	47	389	5	4	6	12	144	46	47	27	
26.休養・くつろぎ	476	65	83	53	23	390	17	10	14	27	159	34	53	15	
27.テレビ・動画・音楽	450	69	84	70	65	386	6	6	15	23	147	56	51	22	
28.読書・新聞	465	79	87	69	45	351	3	6	12	33	161	40	46	19	
29.学習・自己啓発	446	72	86	107	19	418	7	15	3	41	75	36	100	13	
30.その他	467	56	91	71	31	402	7	6	14	37	139	31	45	23	
31.睡眠	464	69	72	55	50	400	9	4	13	37	139	36	27	31	
32.身の回りのこと	449	69	75	66	47	440	1	1	9	14	152	51	37	19	
33.食事	451	69	80	77	40	405	4	2	21	14	135	49	49	11	
34.通勤・通学	459	64	80	67	27	433	18	7	20	152	39	71	17		
35.仕事	432	70	81	67	63	384	7	5	20	25	133	34	63	32	
36.学業	451	76	77	72	62	405	3	9	31	111	61	64	18		

情報Ⅰ

2つずつで
頑張ってみる

	01.睡眠	02.身の回りのこと	03.食事	04.通勤・通学	05.仕事	06.学業
01.睡眠						
02.身の回りのこと	-0.2308					
03.食事	-0.129	0.04927				
04.通勤・通学	-0.3563	0.19467	0.28387			
05.仕事	-0.3274	0.08675	-0.1459	0.02845		
06.学業	-0.3835	-0.1301	-0.2055	-0.0192	-0.0992	
07.家事	0.03418	-0.2337	0.04911	0.06168	-0.1948	0.05732
10.買い物	0.26421	0.02212	0.0845	0.12755	-0.4664	-0.239
11.移動・交通	-0.1549	0.14999	0.02546	-0.107	-0.1786	-0.2163
12.テレビ・動画・音楽	0.35748	-0.1084	0.11338	-0.0222	-0.2739	-0.0922
13.休養・くつろぎ	0.09791	-0.1746	0.04584	-0.1438	0.2955	-0.4069
14.学習・自己啓発	-0.1175	0.37528	-0.0702	0.03141	0.02206	-0.2363
15.趣味・娯楽	0.12724	-0.1379	0.11829	-0.052	-0.3413	-0.2828
16.スポーツ	0.0383	-0.0809	-0.0632	-0.2982	0.01111	-0.1535
20.その他	0.10732	-0.0867	-0.2132	-0.3353	-0.1576	0.01509



相関行列

散布図行列

情報Ⅱ

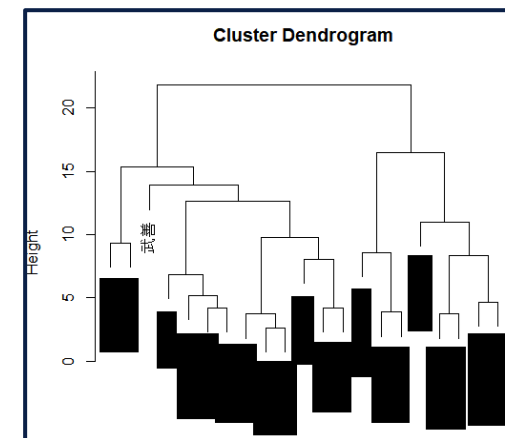
そのまま見る!

主成分分析で、次元を圧縮しよう



文科省「情報Ⅱ」解説動画より

クラスタリングで、グループ化しよう



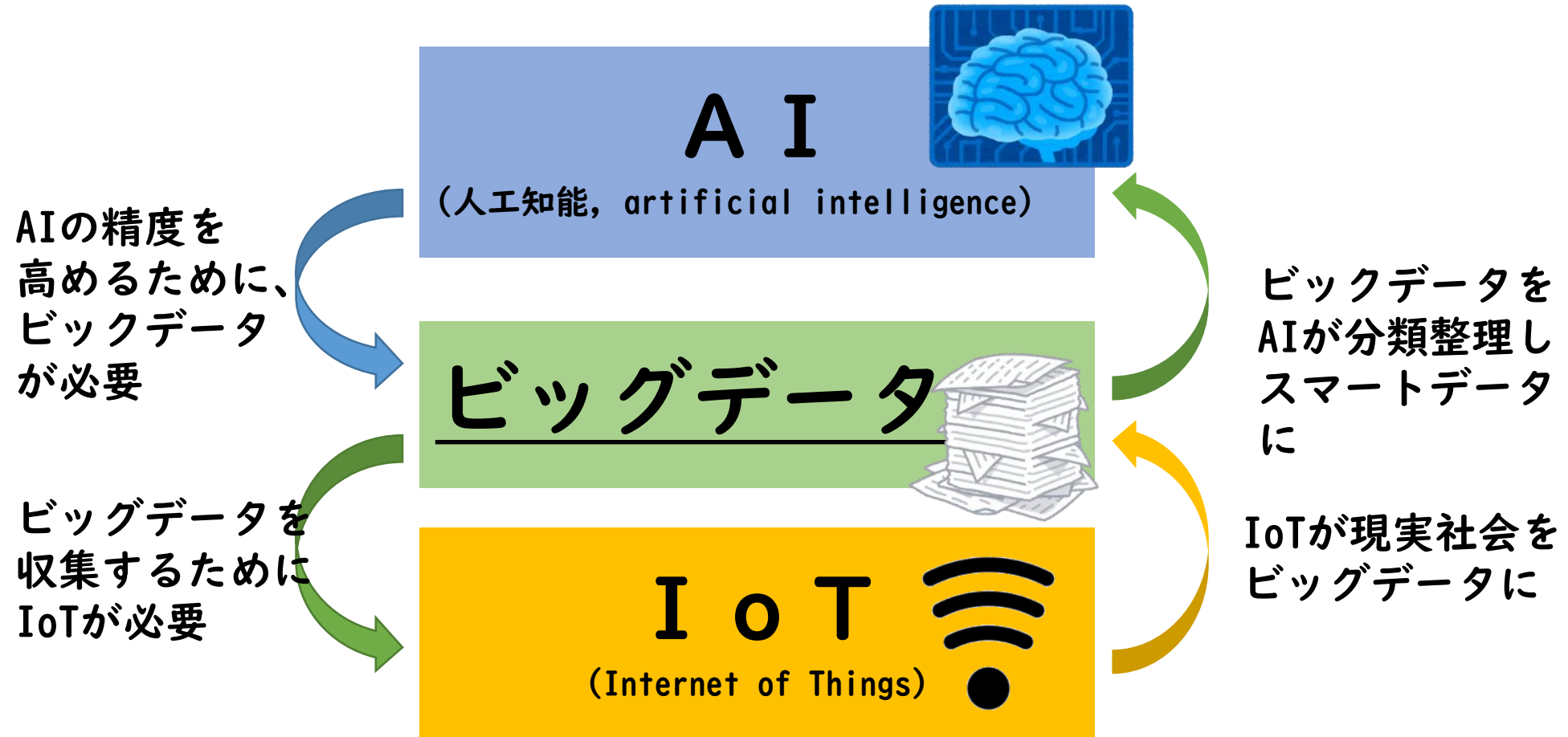
etc...

Agenda.

- 1 情報Ⅱのデータサイエンス(情報Ⅰと比較しながら)
- 2 情報Ⅱの授業演習1 「機械学習」分野
- 3 情報Ⅱの授業演習2 「統計学」分野
- 4 データサイエンスを扱う価値**

①社会の在り方 今後の社会の仕組み(データ駆動型社会)

データは21世紀の**石油**である。



以下Webページの図を参考に作成：

<https://it.impress.co.jp/articles/-/14382>

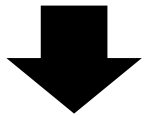
②情報科の要 「試行錯誤」の場面が豊富なのもデータサイエンス

多くの生徒が、勉強のせいで睡眠不足に陥っている！今後、改善したい！
他に考えられる結論は？

結論

Conclusion

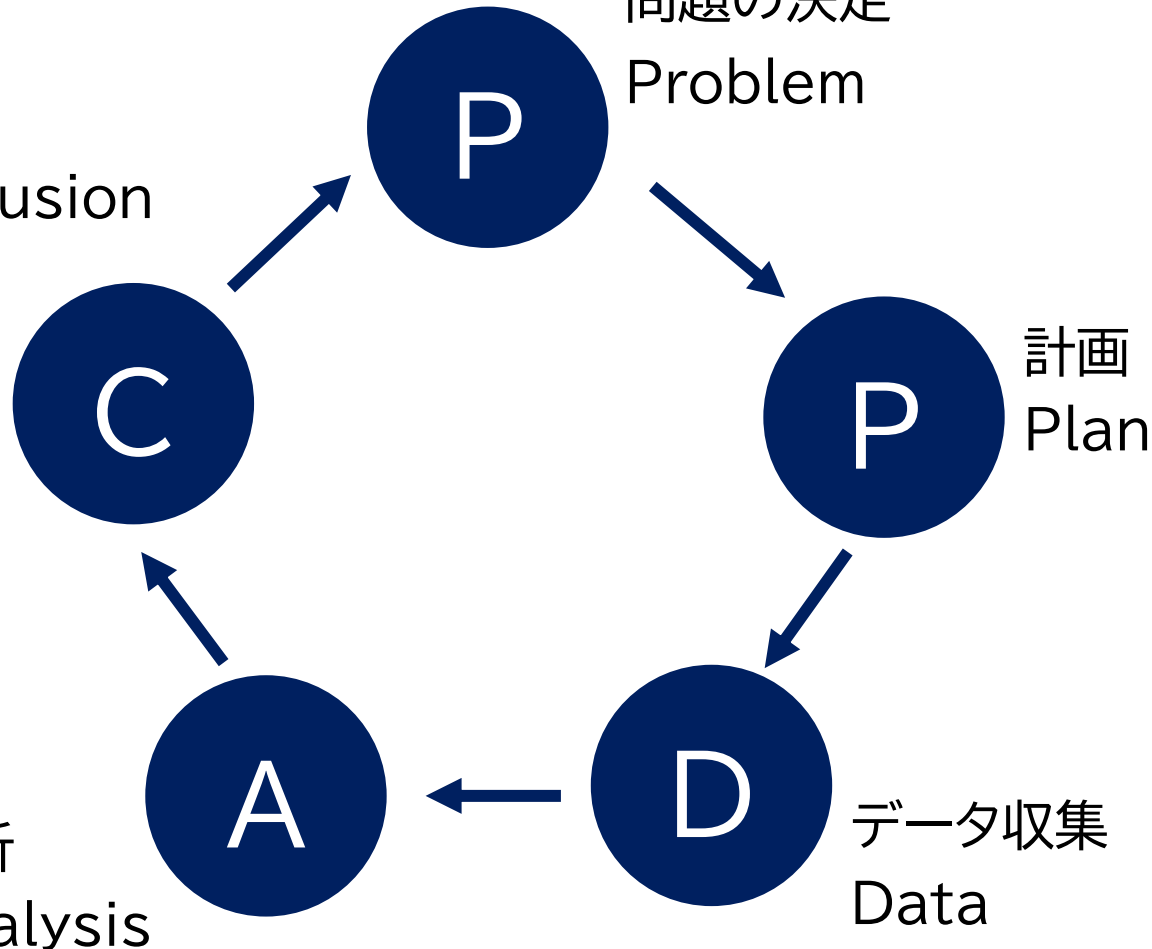
情報Ⅱで視野が大きく広がる！



学習時間が多い人は、睡眠時間が短い！
分析手法を変えてみよう！

分析

Analysis



睡眠時間と学習時間に関係性があるか？
類似するテーマには何がある？

クラスに、アンケート調査をやってみよう

アンケートを実施
アンケートはどのように取る？

③統計(データサイエンス)は人生を豊かにする

「雨乞いをすると雨が降る」……？

	降雨あり	降雨なし
雨乞いあり	A	B
雨乞いなし	C	D

「雨乞いをすると雨が降る」
ためには、
「雨乞いをしないと雨が降ら
ない」
ことも必要

人間

自分の信じたい仮説を支持する
情報を積極的に集めてしまう
(確証バイアス)

→素早い情報処理

コンピュータ・統計

感情に左右されず,数学的に物
事を処理していく。

→綿密な情報処理

統計はパートナー。世の中の見え方を変えてくれる

まとめ

- 情報Ⅱは、案外手を出しやすい。
- 「おっ、今度はこのデータを分析してみたいぞ！」と思えるのは情報Ⅱの方が多いかも。
- ある高校生の言葉
「高校生活で一番感動したことは、主成分分析です」

終

発表資料は全てWebで公開しています

<https://www.hinode.ed.jp/share/takeyoshi/>