好奇心旺盛なデータドリブン思考の育成を目指すデータサイエンスの授業実践

国立大学法人東海国立大学機構

名古屋大学教育学部附属中・高等学校

都丸希和（数学科）　佐藤健太（体育科）

※データドリブン思考
　単なる経験や直感に頼るのではなく、
　客観的なデータを基に意思決定を行うアプローチ

実践校でのDSの位置付け

2-3年生合同

課題研究 STEAM

目的　インターディシプリナリーに課題を捉え、
社会へのつながりを考えて解決へ導く力を養う

方法　少人数ゼミ

- 個人テーマで探究（共同・グループ可）
- 報告会

様々な異分野に
助言や質問

多領域合同報告会

1年生

データサイエンス

量的分析

1年生

アカデミックライティング

質的分析

全ての生徒が客観的に定量評価する視点を持つことを目指す

DSでの協同の意義

複数の視点があると嬉しいとき

- **統計手法の選択**
- **変数の選択**
- **結果の解釈**
- **用いる図の表現方法**

etc..

他者との意見交換

- **自分以外の視点の「ありがたみ」を実感**
- **視点は沢山あると良いことを実感**

DSでの協同を通じて、STEAMでの異分野協同に繋げる

DSの授業構成

STEAMで使いそうな基本的な分析

- グラフ化
- **相関分析**
- 回帰分析
- **平均の差の検定**
- 適合度検定
- 独立性の検定

A. 理論・概念を理解

B. 実際に統計処理を行う

初年度　A → B

2年目　B → A

3年目　？→？

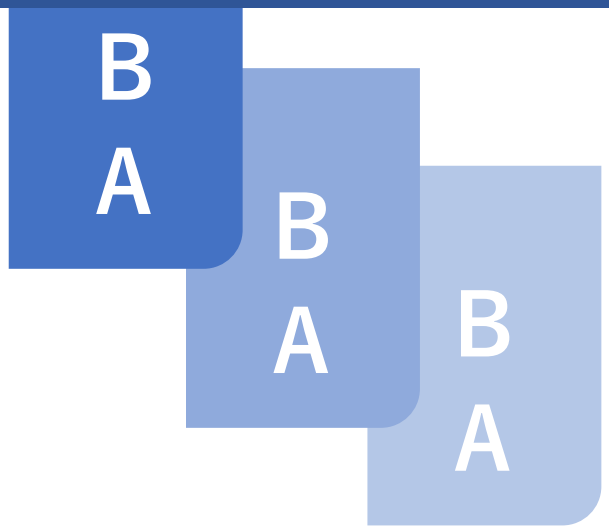
DSの授業構成

3年目のDSは、

a. 使う場面と方法だけ理解

B. 実際に統計処理を行う

A. 理論・概念を詳しく理解



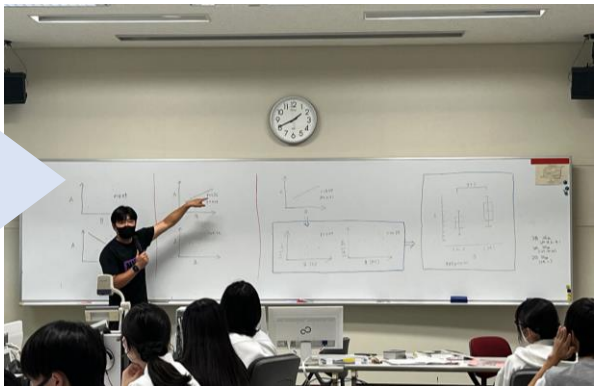
トピックごとにできる限り細かく何度か繰り返す。

DSの授業構成

年間３つのフェーズで構成

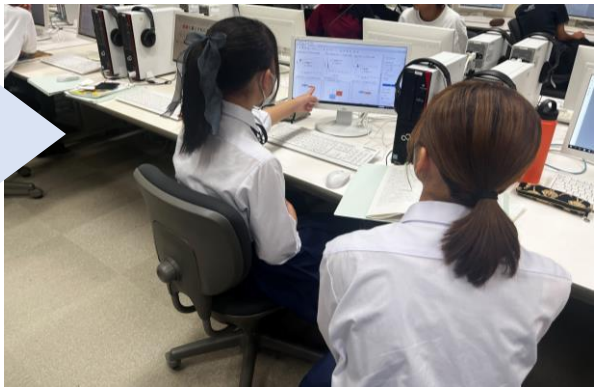
Ⅰ．方法と結果の解釈

演習



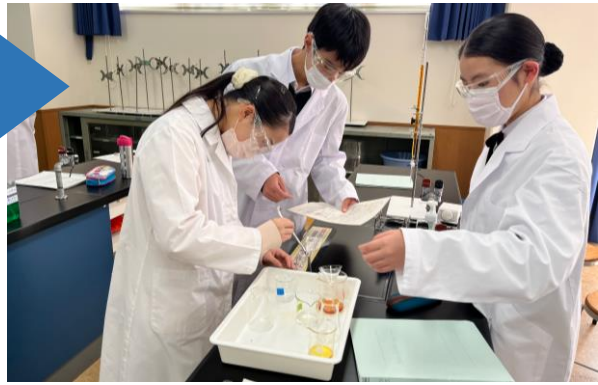
Ⅱ．方法の選択

ポスター・レポート



Ⅲ．データの取得と整理

PBL



データサイエンス（DS）実践者について

体育科

フェーズ1・2
（理論補佐）

数学科

フェーズ1・2・3、数Ⅰ

理科（化学）

フェーズ3（実験）

情報

Excel・Python操作

数学・情報・理科＋他教科で連携

DS前期指導計画（フェーズⅠ・Ⅱ）

回		内　容
1	理論	基礎　データの種類・構造
2		基礎　統計図表・代表値
3		基礎　散布度
4	実践	演習　相関関係
5		基礎　回帰分析・時系列分析
6		基礎　仮設検定
7		演習　t検定
8		基礎　クロス集計表・適合度検定
9		応用　データ収集・正規分布
10		応用　統計ポスター作成
夏		応用　統計ポスター作成・提出
11		基礎　レポート作成時の注意
12		基礎　分析・レポート作成
13		応用　分析・レポート作成２回目
14		応用　分析・レポート完成・まとめ

情報の授業で
エクセルの基礎演習

**情報・数学と
重複する部分**

数学Ⅰ・Bの授業で
問題演習

Let’s　TRY!!

次のデータ（変数）を２種類に分けよう！

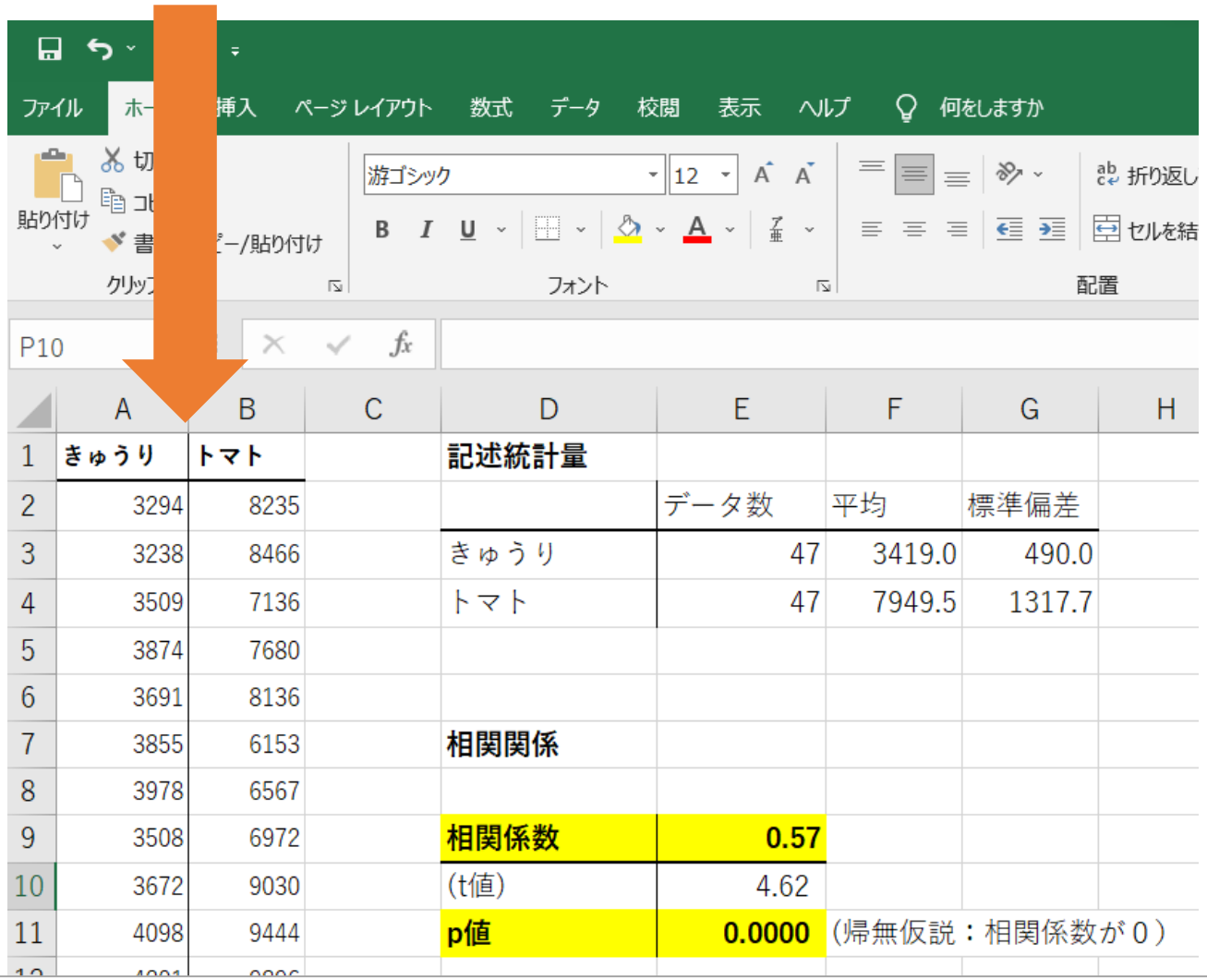
1　クラス（A、B、C組）
2　出席番号
3　身長
4　年齢
5　色（赤、黄、青･･･）
6　かけっこの順位
7　走り幅跳びの記録
8　国語のテストの点数
9　数学の成績（５段階）

他者との意見交換

フェーズI：オープンソースを用いた演習

エクセルで相関分析をする（ 方法）

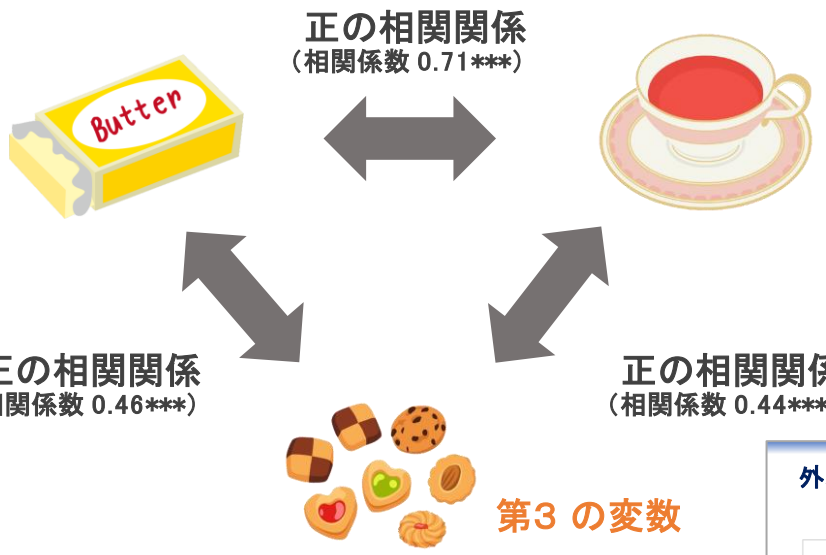
◆ 相関関係を調べたいペアをA列とB列にペースト



	A	B	C	D	E	F	G	H
1	きゅうり	トマト		記述統計量				
2	3294	8235			データ数	平均	標準偏差	
3	3238	8466		きゅうり	47	3419.0	490.0	
4	3509	7136		トマト	47	7949.5	1317.7	
5	3874	7680						
6	3691	8136						
7	3855	6153		相関関係				
8	3978	6567						
9	3508	6972		相関係数	0.57			
10	3672	9030		(t値)	4.62			
11	4098	9444		p値	0.0000	(帰無仮説：相関係数が０)		

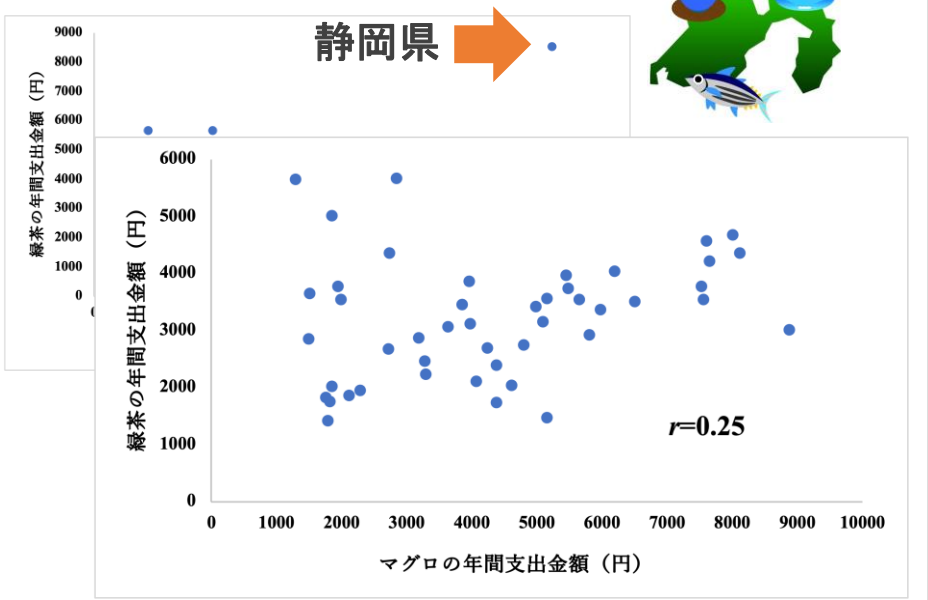
フェーズI：相関係数の読み取り

相関関係と因果関係



相関関係は2 種類の関係性を調べるもの！

外れ値の影響

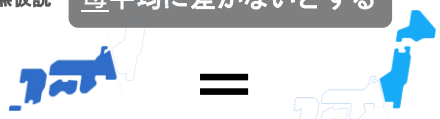


他者との意見交換

フェーズI： *t* 検定の考え方

仮説検定の考え方　t 検定の仕組み

帰無仮説　母平均に差がないとする



標本西　　標本東

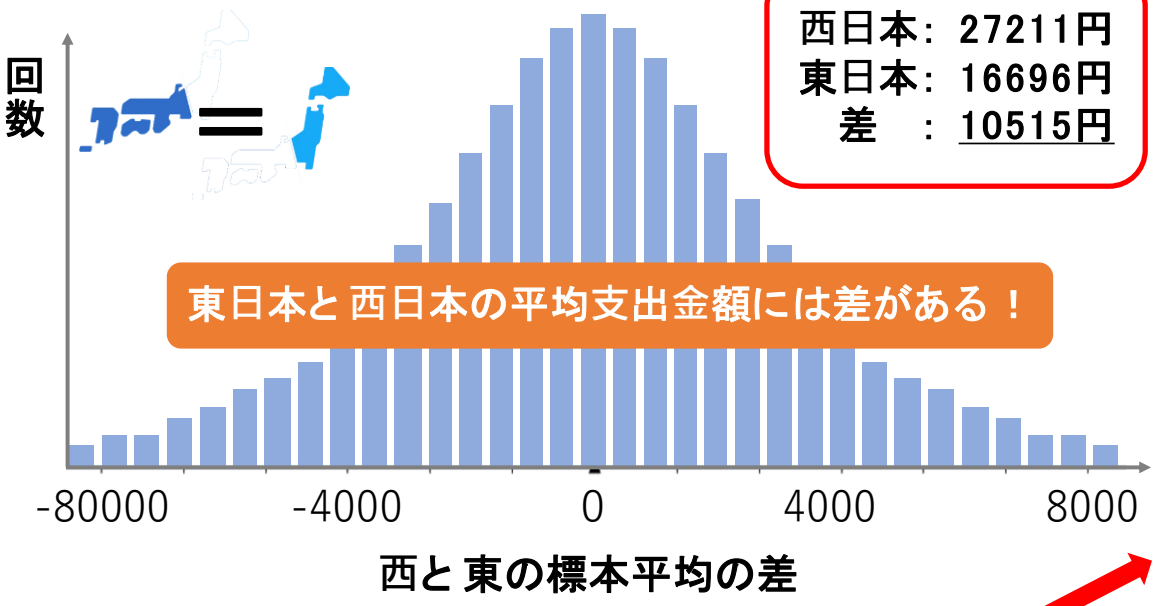
（標本西の平均）－（標本東の平均）を記録

1000回分記録する

仮説検定の考え方　標本分布

今回の標本の平均

西日本：27211円
東日本：16696円
差　：10515円



東日本と西日本の平均支出金額には差がある！

1000個の標本で10515円の差が出たのは
1 個もなかった！

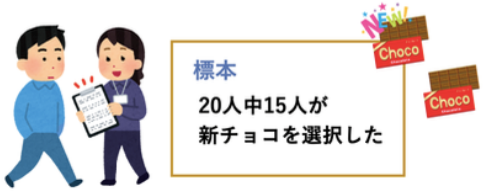
他者との意見交換

フェーズIをまとめたテキストの作成

第 2 部　仮説検定の考え方

母集団の推測

チョコレートメーカーが新商品について、アンケートを取りました。20人に旧商品と新商品の2種類を食べてもらい、どちらが美味しいかを質問しました。このとき15名が新チョコを美味しいといった場合、新チョコは美味しくなったと判断して良いでしょうか。



母集団（消費者全体）でも同じような結果が得られる？

まず、帰無仮説を考えましょう。ここでの対立仮説（言いたい事）は、「新チョコが美味しくなった」です。ということは、帰無仮説は「新チョコと旧チョコで味は変わらない」です。具体的には、美味しさが同じということで「消費者全体の意見は半々に分かれる」という仮説になります。

帰無仮説　**新旧のチョコで味が変わらない場合
消費者全体の意見は半々になる**

半々ということで、コイントスに置き換えて考えてみます。コイントスを20回やり、表が何回出たかを記録します。これを1セットとして、100セット試行します。その結果、以下のようになりました。

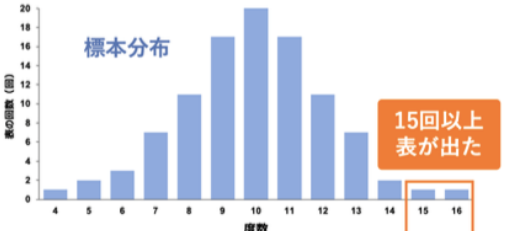
コイントスでシミュレーションした結果

表の回数	4	5	6	7	8	9	10	11	12	13	14	15	16	計
度数	1	2	3	7	11	17	20	17	11	7	2	1	1	100

標本分布

各標本の結果（いまの場合、コイントスでシミュレーションした結果）をグラフで表したものを、標本分布といいます。

15人以上が「新チョコ」を選ぶ確率は、100セットの中で15回以上表が出たセット数の確率と等しくなります。今回、表が15回以上であったのは、100セットのうち2セットでした。

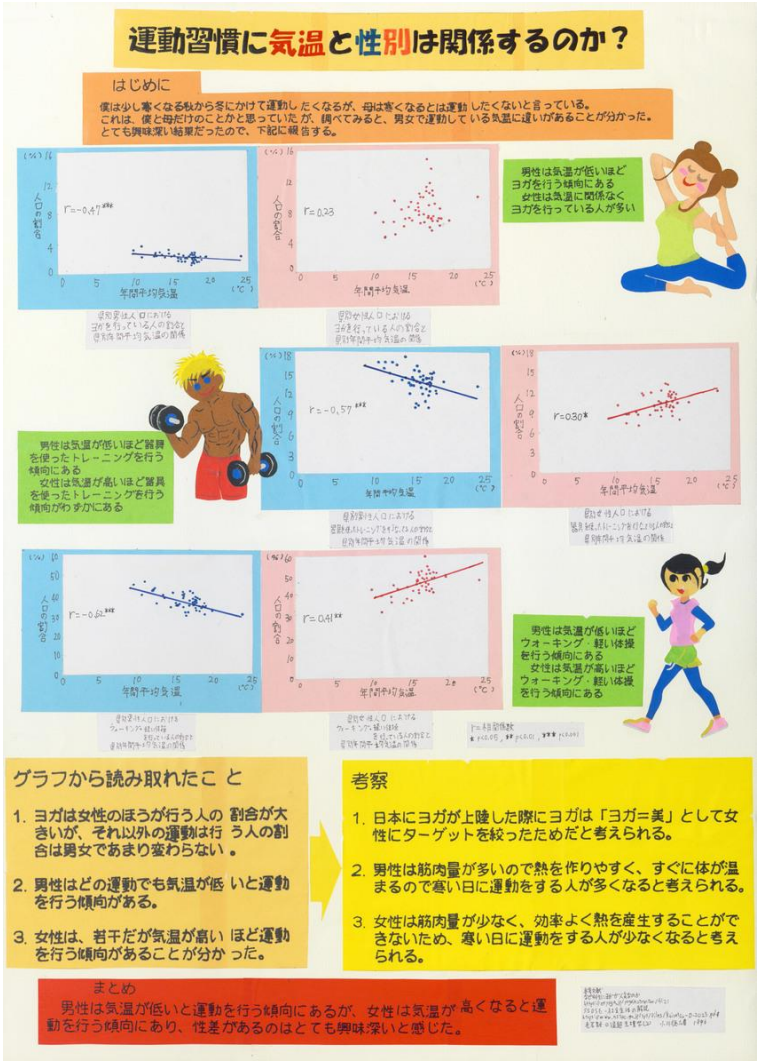


表の回数	4	5	6	7	8	9	10	11	12	13	14	15	16	計
度数	1	2	3	7	11	17	20	17	11	7	2	1	1	100

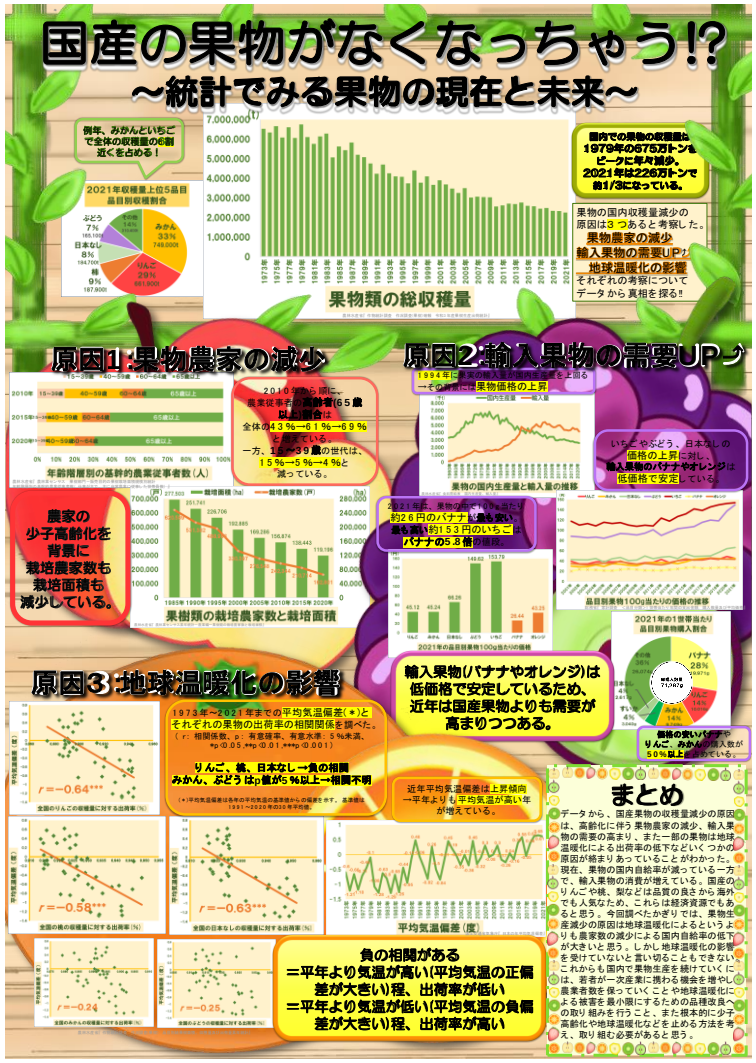
表が出た回数は100セット中2セットで確率0.02

※監修：名古屋大学大学院教育発達科学研究科 石井秀宗教授

フェーズII：統計ポスターの作成（愛知県グラフコンクールへの参加）



第５部（高校生以上手書きの部）
愛知県金賞作品



PC統計グラフの部（小学生以上）
愛知県金賞・全国特選作品

フェーズII：ポスターのソースをもとにしたレポートの作成

データサイエンス（前期）まとめレポート **2024-03-01**

高校１年（　）組（　）番氏名（　　）

1. 仮説

既存の電子レンジは消費者、特に所得の高い人のニーズに合っていない。

2. 分析方法

【方法A】

電子レンジの所有率を調べた。

【方法B】

電子レンジの機能を使いこなせていると感じている人の割合を調べた。

【方法C】

キッチン家電を使う際に意識することの項目別の割合を調べた。

【方法D】

47 都道府県の一人当たりの所得と調理食品への年間支出金額との関係を調べるため、相関分析を行った。

【方法E】

47 都道府県の一人当たりの所得と趣味としての料理・お菓子作りにかける時間との関係を調べるため、相関分析を行った。

【方法F】

47 都道府県の一人当たりの所得の平均（3003830 円）より高い群と低い群に分けて、2 群間の調理食品の年間支出金額に対する等分散を仮定しない 2 標本による t 検定を行った。

【方法G】

47 都道府県の一人当たりの所得の平均（3003830 円）より高い群と低い群に分けて、2 群間の趣味としての料理・お菓子作りにかける時間に対する等分散を仮定しない 2 標本による t 検定を行った。

3. 結果

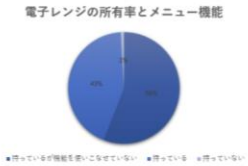
表 1. 記述統計量（　日本における所得と調理食品　）

調理食品	136911.48	±	12273.37
一人当たり県民所	2951.41	±	302.99
趣味としての料理・お菓子作り	17.79	±	8.10

値は平均値±標準偏差で示す

【結果A】

電子レンジの所有率を調べたところ、９９％以上の人が持っていると答えた。

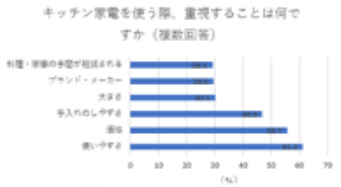


【結果B】

電子レンジのメニュー機能を使いこなせていると感じているかを調べたところ、使いこなせていると答えた人は４３．４％なのに比べ、使いこなせていないと答えた人は５６．６％になった。

【結果C】

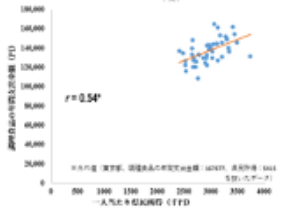
キッチン家電を使うときに重視することの項目別の割合（複数回答可）を調べたところ、「使いやすさ」の項目は 6 割以上の人に重視されており、一番重視する人が多かった。



【結果D】

一人当たり県民所得と調理食品の年間消費支出との相関係数を調べた。その際、県民所得 1 位の東京都と 2 位の愛知県の差が 1617 円になったため、東京都を外れ値としてほかの 46 道府県の相関係数を調べたところ、相関係数は 0.54、有意確率は 0.0001 であり、有意な正の相関がみられた。（図のアスタリスクは正しくは 3 個）

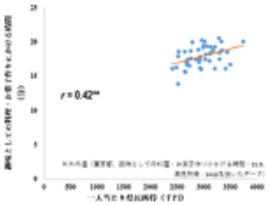
*** p < 0.001



【結果E】

一人当たり県民所得と趣味としての料理・お菓子作りにかける時間との相関係数を調べた。その際、県民所得 1 位の東京都と 2 位の愛知県の差が 1617 円になったため、東京都を外れ値としてほかの 46 道府県の相関係数を調べたところ、相関係数は 0.42、有意確率は 0.0036 であり、有意な正の相関がみられた。

**p < 0.01



【結果F】

一人当たり県民所得が低い地域（N=26）と高い地域（N=21）の調理食品の年間支出金額の違いを比較したところ、調理食品の年間支出金額は一人当たり県民所得が高い地域が低い地域に比べ有意に高かった。（図１）

【結果G】

一人当たり県民所得が低い地域（N=26）と高い地域（N=21）の趣味としての料理・お菓子作りにかける時間の違いを比較したところ、趣味としての料理・お菓子作りにかける時間は一人当たり県民所得が高い地域が低い地域に比べ有意に高かった。（図２）

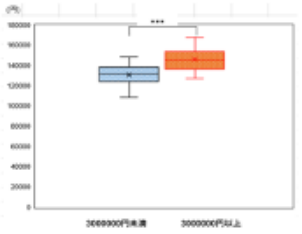


図１　**p < 0.001

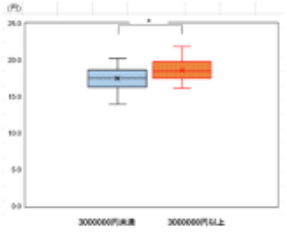


図２　*p < 0.05

第6 問　これあなたならどう 書きますか？

結果

即席麺の年間支出金額には、群間差は見られなかった。

これはそもそも即席麺の需要が低いことが影響していると考えられる。

➡ 削除

結果から 予想される 背景や因果関係などは考察へ。ただし 、即席麺の需要が低いという 根拠を 記載（ 先行研究） 。

“寒い地域で温かいラーメンが好まれる”

一般論は根拠必要なし で良い場合が多い

他者との意見交換

DS後期指導計画（フェーズⅢ）

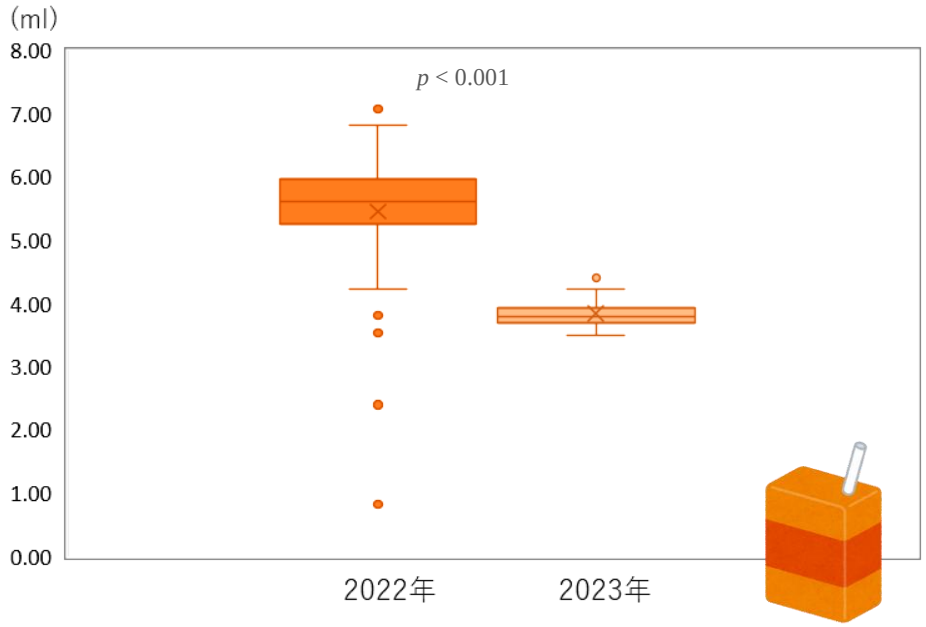
回		内　容
1	準備	後期オリエンテーション
2		滴定原理・器具の説明
3		濃度既知のビタミンCの定量滴定
4		プレ実験（オレジュース対グレープフルーツジュース）
5	実践	分析・グループ自由テーマ検討
6		PBL研究テーマ発表会・実験計画
7		PBL実験（ビタミンCの滴定）①
8		PBL実験（ビタミンCの滴定）②
9		PBL実験レポート作成
10		PBL実験レポート作成・再実験計画
夏		PBL再実験（ビタミンCの滴定）
11		PBL実験レポート提出
12		PBL実験レポート発表会
13		相関・重回帰分析
14		まとめ・アンケート

実験環境の整備

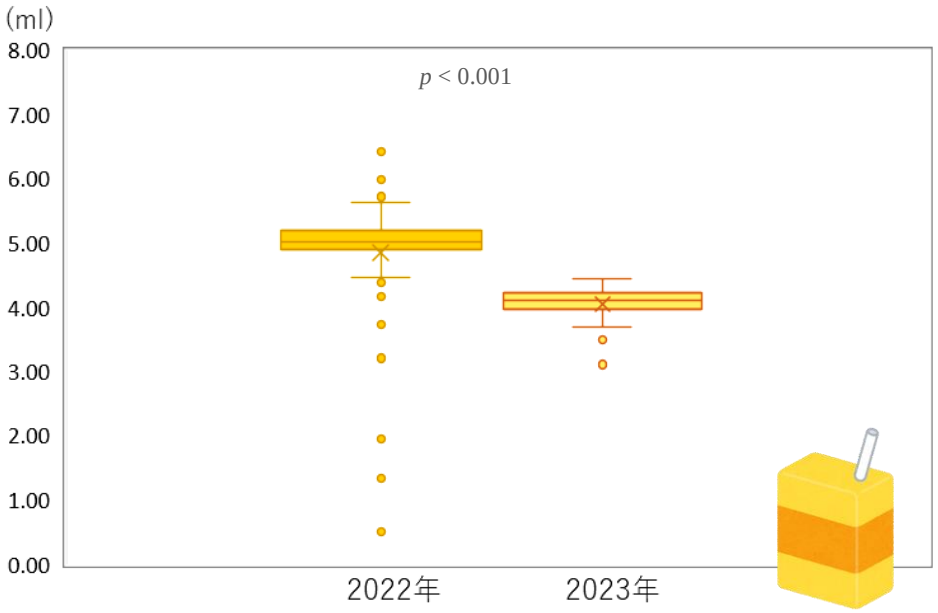
実験による
データを生成する経験

フェーズⅢ：導入問題としての予備実験（2022年と2023年の比較）

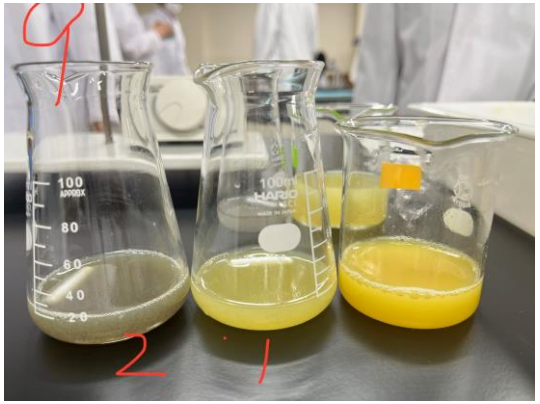
オレンジジュースとグレープフルーツジュースはビタミンCの含有量に違いがある？



オレンジジュース



グレープフルーツジュース



他者との意見交換

フェーズⅢ：オリジナル試料のテーマ発表会

実験・分析に入るまでのイメージ

① ビタミンC含量を調べる資料を選ぶ
② ①を選んだ理由
③ 仮説を立てる
④ 具体的な予想をする
⑤ ④のように予想した理由

実現可能性を意識

⑥ データの収集と分析の方法
⑦ 試料収集の際に気を付けること
⑧ 実験やデータ収集の際に気を付けること

他者との意見交換

フェーズⅢ：生徒が選んだテーマ例

1班　🍊🍎🍇１００％果汁ゼリーの味によって、ビタミンCの含有量は異なるのか？（同一メーカー、袋）

2班　💴値段とビタミンCの量は比例するのか 試料：オロナミンCとオロナミンCゴールド

3班　🍅小さいトマトの方が大きい🍅よりビタミンCを多く含んでいる（熊本県産、愛知県産（２回目～））

4班　🥝ビタミンC含量はゆでることや電子レンジであたためることで変化するのか（キウイ、福岡県産）

5班　🍋ゆでることでレモンのビタミンC含有量は変化するのか（チリ産とアメリカ産（３回目以降））

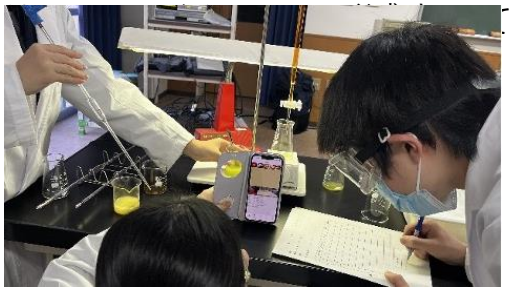
6班　🥝キウイフルーツにおいて、緑色の部分は白色の部分に比べてビタミンCの含有量が多い（NZ産）

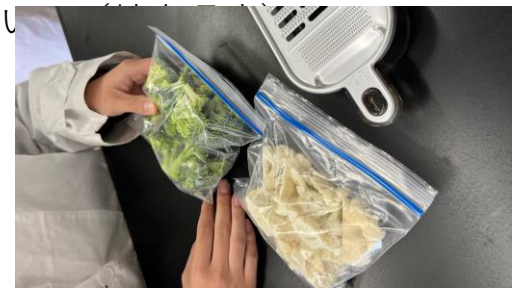
7班　🍊同じ日本産蜜柑でも品種が違えばビタミンCの量が変わるのか（静岡県産、熊本県産）

8班　🍅トマトの色によってビタミンCの含まれている量が違うのか（愛知県産）

9班　🍅柿の種類によってビタミンCの含有量は異なるのか（スーパー、Mさんの家）

10班　🥬ゆでた後のホウレンソウは、ゆでる前に





フェーズⅢ：ビタミンCの滴定実験レポートの作成

高校１年生　データサイエンス　　ビタミンＣ測定レポート　　２０２４．１
　１年（　Ｂ　）組（　　）番　名前（　　）
　（　８　）班　　班員の名前（　　）

１．研究テーマ・試料・仮説

ブロッコリーとカリフラワーでは含まれるビタミンＣの量に違いがある。
また、ブロッコリーのほうがカリフラワーよりもビタミンＣ含有量が多い。

① ビタミンＣ含量を調べる試料
・ブロッコリー（香川県産）（マックスバリュ本山店）（切り口にラップをかけ、ビニール袋の中に入れ冷凍保存）
・カリフラワー（三重県産）（マックスバリュ本山店）（ビニール袋の中に入れ冷凍保存）

② ①を選んだ理由
ブロッコリーがビタミンＣの多い食べ物ランキング四位（https://www.kusurinomadoguchi.com）であることに意外性を感じ、同時に食感や見た目が似ており、
「ブロッコリーの突然変異で花蕾が白化したものといわれ」（https://www.takii.co.jp）るカリフラワーとは違いがあるのか、私たちによりなじみのあるブロッコリーを食べることは損になっていないのか気になったため。

③ 仮説とその理由
　２０２０年作物統計調査では、カリフラワーの生産量２１０００トン、出荷量１８０００トンなのに対し、ブロッコリーは生産量１７４５００トン、出荷量１５８２００トンだった。また農林水産省のホームページに示された「身近な作物」の中に、カリフラワーの姿を見ることはできなかった。その明らかな格差からブロッコリーのほうがカリフラワーより選ばれていると捉えることができると考える。これ　は栄養価、その中でもビタミンＣの含有量の多いブロッコリーを人々が求めているからではないかと考えた。

「アスコルビン酸生合成は光によって厳密に調節されており、光強度が高ければ高いほど活性化され、逆に光のない条件（夜間）では速やかに不活性化されます。」（https://shimane-univ-biochemistry.jp）とあるが、ブロッコリーは熱を吸収しやすい緑色であるのに対しカリフラワーは白色であることから、ブロッコリーのほうがビタミンＣを多く生合成し、蓄えていると考えた。また、「ビタミンＣは、植物が強い光にさらされると、細胞質のミトコンドリアで作られて、同じ細胞内の葉緑体へ運ばれ、過剰な光エネルギーを熱として逃して、葉の光障害を防ぐ。」（https://scienceportal.jst.go.jp/newsflash/20150107_01）とされており、花も緑色であることからブロッコリーにはカリフラワーと比べ葉緑体も多く含まれると考え、販売されている一房に葉緑体が含まれる部分（ビタミンＣを蓄える場所）が多い分一房当たりの含量も多くなるのではないかと考えた。
そこでこの実験に臨む上での仮説を、
　「ブロッコリーとカリフラワーのビタミンＣ含量には有意差があり、カリフラワーよりブロッコリーのほうがビタミンＣを多く含む。」とする。

２．　実験方法

Ａブロッコリー

冷凍したブロッコリー一株をブロッコリーが潰れないようにお湯につけ、包丁や手で房に分けて、ジューサーでさらに細かくする。
その後、不織布で細かくしたブロッコリーを包み、手でビーカーに液体を絞り出す。

Ｂカリフラワー

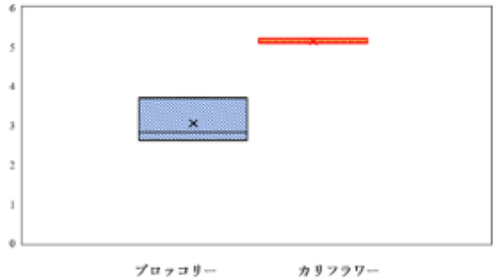
冷凍したブロッコリー一株をブロッコリーが潰れないようにお湯につけ、包丁や手で房に分けて、ジューサーでさらに細かくする。その後、不織布で細かくしたカリフラワーを包み、手でビーカーに液体を絞り出す。
① ヨウ素液をビュレットに入れ、目盛りをよむ。
② 資料Ａをホールピペットで１０ｍｌ測りとり、コニカルビーカーに入れる。
③ ②にデンプン溶液を二滴加える。
④ スターラーで撹拌しながら、ビュレットからヨウ素液を滴下する。
⑤ 色が紫色っぽく変化し、戻らなくなった時点を反応の終点として、ビュレットの目盛りを読む。
資料Ｂにおいても①～⑤を繰り返しＡ、Ｂ各３回(12 月 13 日)、各４回（1 月 31 日）の酸化還元滴定を行った。

３．結果（元データとデータ処理方法も書くこと）
表１　ブロッコリーとカリフラワーの滴定値（12 月 13 日）
　　数値は加えたヨウ素液の体積(ml)
　　ヨウ素液濃度はすべて 0.0040mol/L
記述統計量 n=3

	ブロッコリー	カリフラワー
１回目	3.70	5.21
２回目	2.82	5.09
３回目	2.63	5.10
	平均値	標準偏差
ブロッコリー	3.05	0.57
カリフラワー	5.13	0.07

ｐ値 0.02288＜0.05

フェーズⅢ：ビタミンCの滴定実験レポートの作成



↑ヨウ素液の体積（ml）

表２ ブロッコリーとカリフラワーの滴定値（１月 31 日）

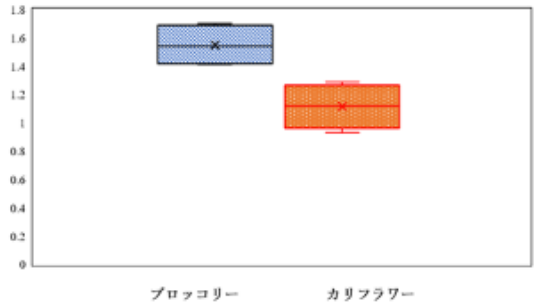
数値は加えたヨウ素液の体積(ml)

ヨウ素液濃度はすべて 0.0038mol/L

記述統計量 n=4

	ブロッコリー	カリフラワー
１回目	1.65	1.19
２回目	1.70	1.29
３回目	1.41	0.93
４回目	1.43	1.05
	平均値	標準偏差
ブロッコリー	1.55	0.15
カリフラワー	1.12	0.16

p 値 0.00727＜0.01



↑ヨウ素液の体積（ml）

４．考察　　ここまでの実験でわかったこと

今回、各７回ずつの実験を行った。前半の３回と後半の４回にわけて考察する。

まず、前半である。ブロッコリー１リットルにおけるビタミンｃの量は、

0.0040×3.05=c×10

c=0.00120mol/L

0.00120×176=0.2112…

1L に 0.211mg 含まれる

カリフラワー1 リットルにおけるビタミンｃの量は、

0.0040×5.13=c×10

c=0.00205mol/L

0.00205×176=0.3608…

1L に 0.361mg 含まれる

次に、後半である。ブロッコリー１リットルにおけるビタミンｃの量は、

0.0038×1.55=c×10

c=0.000589mol/L

0.000589×176=0.1036…

1L に 0.104mg 含まれる

カリフラワー1 リットルにおけるビタミンｃの量は、

0.0038×1.12=c×10

c=0.0004256mol/L

0.00205×176=0.07490…

1L に 0.075mg 含まれる

前半のｔ検定のｐ値が＊ではあるが、前半ではカリフラワーのビタミンＣ含有量がブロッコリーのビタミンＣ含有量よりも多かったにもかかわらず、後半ではブロッコリーのビタミンＣ含有量がカリフラワーのビタミンＣ含有量よりも多くなっている。その原因を考察したい。

考えられる原因は主に３つある。

１つ目は、ブロッコリーとカリフラワーのビタミンＣが減る量の違いである。冷凍はしていたが１回目の実験と２回目の実験との期間が長く、ビタミンＣが減ってしまった。具体的には、購入した日がブロッコリーの方がカリフラワーより１週間ほど早かった。

２つ目は、でんぷん反応で色が変わったと判定する基準の違いである。これも、１回目の実験と２回目の実験との期間が長く、１回目で色が変わったと判断したときに写真を撮っていなかったり、実験をする人が異なっていたりした。

３つ目は、ブロッコリーとカリフラワーを解凍する方法の違いである。実験方法で、どちらも袋に入れた状態でお湯につけたと述べたが、実際には、１回目ではブロッコリーのみ太陽光に当て、２回目ではカリフラワーのみ袋に入れずに直接お湯につけたことがあった。

５．再実験の計画

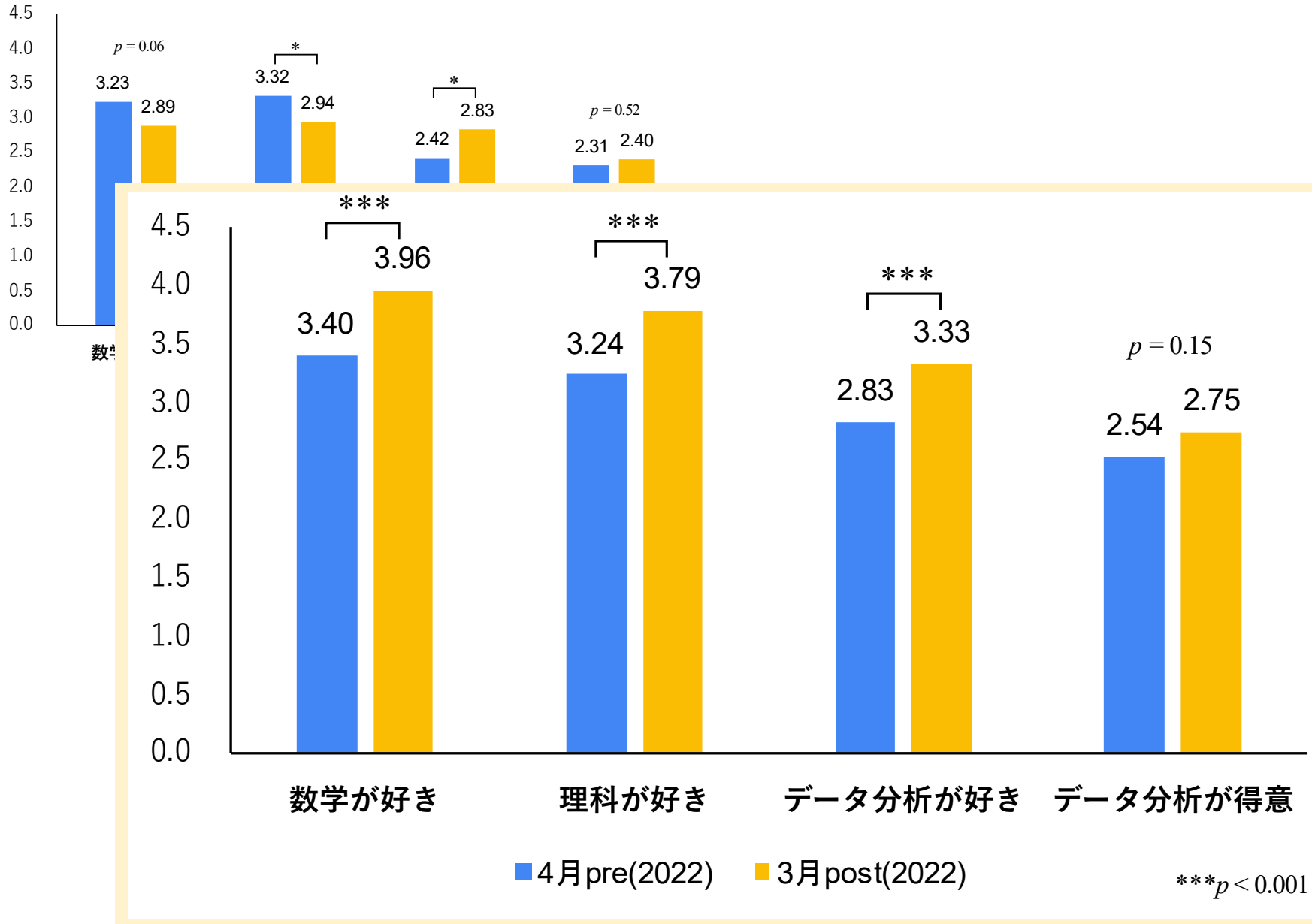
気をつけることはデンプン反応で色が変わったとする基準を揃えることである。特にブロッコリーは徐々に茶色く濁っていくが、これはヨウ素液の茶色なのか、ブロッコリーの緑色とデンプン反応の青紫色が混じった色なのかが分かりづらい。次の実験では１回目の写真を撮り、それを基準として目盛を読むようにする。

アンケート結果（2022年）

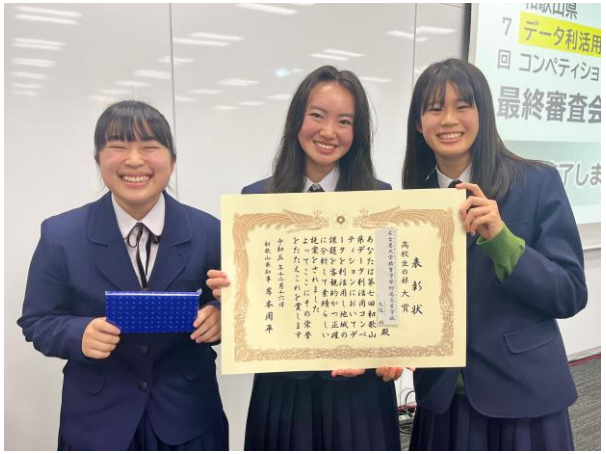
質問項目	4月	3月	
現在「数学」は好きですか	3.40	3.96	***
現在「理科」は好きですか	3.24	3.79	**
現在「統計」「データ分析」は好きですか	2.83	3.33	***
現在「統計」「データ分析」は得意ですか	2.54	2.74	
グラフや表を適切に読むことが出来ますか	3.43	3.26	
実験や観察，アンケートなど自分でデータを収集することが出来ますか	3.06	3.23	
データの種類からどんな分析方法が適切かを判断できますか	2.85	3.20	*
先行研究などを調査して，課題を発見することが出来ますか	2.83	3.31	***
データを加工・整理することが出来ますか	2.88	3.14	
データ解析から結論を導くことが出来ますか	2.97	3.39	**
複数のデータを組み合わせて結論を導くことが出来ますか	3.09	3.30	
自分の研究を客観的に他者に説明することが出来ますか	3.26	3.41	
他の人の研究にアドバイスを出来ますか	2.88	2.83	
他の人の意見を柔軟に取り入れて，考えを修正することが出来ますか	3.59	3.80	*
あなた自身にとって，DSの授業はどの程度必要性を感じますか	2.87	4.25	***

vs.4月（n=120，$^{*}p$<.05,$^{**}p$<.01,$^{***}p$<.001）

アンケート結果（2021年と2022年の比較）



データコンペティションなど発表の機会



和歌山県データ利活用コンペティション（大賞・協賛企業賞）

ユニセフシンポジウムでの発表

WWL高校生フォーラム（優秀ポスター賞）

WWL生徒研究発表会（最優秀ポスター賞）

SDGs将来世代創造フォーラムでの発表