

SSDSEを使った データ活用能力の育成

～統計データ分析コンペティションへの応募を目指して～

愛知県立一宮高等学校 鈴木淳子

発表の流れ

<前半>

SSDSEを活用した
授業の取り組み

<後半>

統計数理賞受賞論文の紹介

勤務校紹介

愛知県立一宮高等学校

普通科

8クラス×3学年

ファッション
創造科

1クラス×3学年



SSH

データサイエンス教育

重点目標の1つ

情報Ⅰ 単元:データの活用 授業の流れ

① 実習:表計算
ソフトの基本

② 実習
e-Stat

③ 実習:
SSDSE

④ 夏課題
レポート作成

④' データ
分析コンペ応募

① 表計算ソフトに慣れる

① 実習:表計算
ソフトの基本

② 実習
e-Stat

③ 実習:
SSDSE

④ 夏課題
レポート作成

④' データ
分析コンペ応募

目的は、合計や平均、四則演算がささっとできるように、
というだけのことだが、意外に時間と手間がかかる

- ・そもそもコンピュータの扱いに慣れていない
→ ダブルクリック・右クリック・ドラッグ・半角/全角....
- ・表計算ソフトに慣れていない
→ 数式を入れるべきところに数値を入れてしまう など

高校入学までに
マスターしてきて
ほしい・・・

② 実習:e-Stat

① 実習:表計算
ソフトの基本

② 実習
e-Stat

③ 実習:
SSDSE

④ 夏課題
レポート作成

④' データ
分析コンペ応募

e-Stat 統計で見る日本
政府統計の総合窓口

e-Statは、日本の統計が閲覧できる政府統計ポータルサイトです

統計データを探す 統計データの活用 統計データの高度利用 統計関連情報 リンク集

●統計データを探す (政府統計の調査結果を探します)

その他の絞込

すべて 政府統計一覧の中から探します

分野 17の統計分野から探します

組織 統計を作成した府省等から探します

キーワード検索: 例: 国勢調査 **検索**

●統計データを活用する

グラフ 主要指標をグラフで表示 (統計ダッシュボード)

時系列表 主要指標を時系列表で表示 (統計ダッシュボード)

地図 地図上に統計データを表示

地域 都道府県、市区町村の主要データを表示

○都道府県データ ○市区町村データ

データ表示

地域ランキング

② 実習:e-Stat

① 実習:表計算
ソフトの基本

② 実習
e-Stat

③ 実習:
SSDSE

④ 夏課題
レポート作成

④' データ
分析コンペ応募

ミッション「愛知県の市町村で外国人居住者の多い市を探せ！」

→ e-Statからデータを取り出し、
表計算ソフトでデータを加工して考える

「外国人が多い地域でゴミ出しトラブルが多い、という話を聞いてデータで真偽を確かめたいとします。ここでいう『多い』とは？単に数の多い少ないの比較でよいですか？」

人口も必要？

表計算ソフトで
割合を算出？

e-Stat に触れた後、SSDSE へ



e-Statは自由度が高すぎるため、初心者が扱いやすいデータが得られる保証はない

- 欠損値
- 複数年度・複数項目で取得した場合にうまく取り出せない など

※ とはいえ、e-Statは多種多様なデータが取得できるため、
独自でテーマ設定をして分析を行う際に使えるように簡単な実習は行う

③ 実習:SSDSE

① 実習:表計算
ソフトの基本

② 実習
e-Stat

③ 実習:
SSDSE

④ 夏課題
レポート作成

④' データ
分析コンペ応募

【実習】

R5 共通テスト 数学Ⅰ・数学Aの問題を使って
データを可視化してみよう

授業スライド
抜粋

① 【SSDSE】うなぎと焼き鳥の支出金額の可視化



R5 共通テスト 数学Ⅰ・数学Aの問題を使って
データの可視化をしてみよう

数学I・数学A

第2問 (必答問題) (配点 30)

〔1〕 太郎さんは、総務省が公表している2020年の家計調査の結果を用いて、地域による食文化の違いについて考えている。家計調査における調査地点は、都道府県庁所在市および政令指定都市(都道府県庁所在市を除く)であり、合計52市である。家計調査の結果の中でも、スーパーマーケットなどで販売されている調理食品の「二人以上の世帯の1世帯当たり年間支出金額(以下、支出金額、単位は円)」を分析することにした。以下においては、52市の調理食品の支出金額をデータとして用いる。

太郎さんは調理食品として、最初にうなぎのかば焼き(以下、かば焼き)に着目し、図1のように52市におけるかば焼きの支出金額のヒストグラムを作成した。ただし、ヒストグラムの各階級の区間は、左側の数値を含み、右側の数値を含まない。

なお、以下の図や表については、総務省のWebページをもとに作成している。

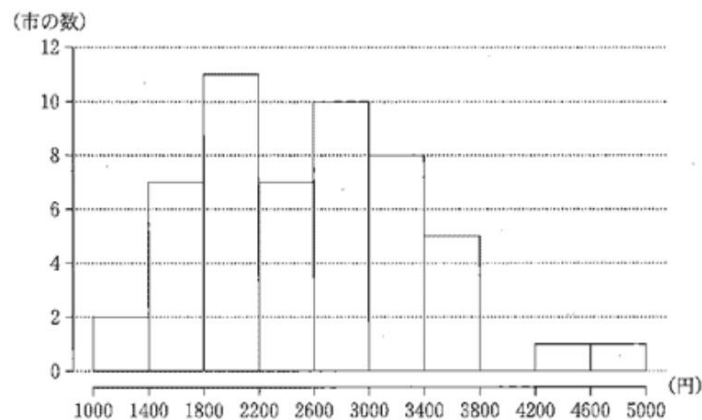


図1 かば焼きの支出金額のヒストグラム

うなぎのかば焼きと焼き鳥の支出金額について
総務省の2020年家計調査から分析。
調査対象は52市。

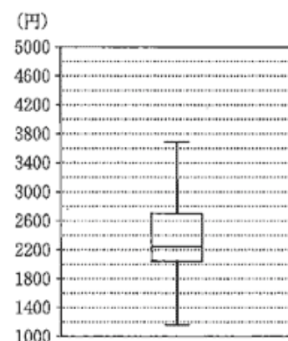


図2 地域Eにおけるかば焼きの支出金額の箱ひげ図

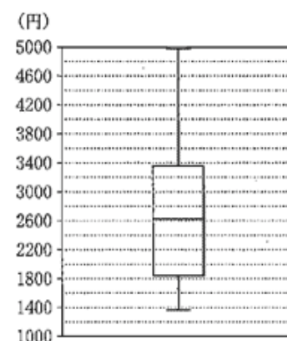


図3 地域Wにおけるかば焼きの支出金額の箱ひげ図

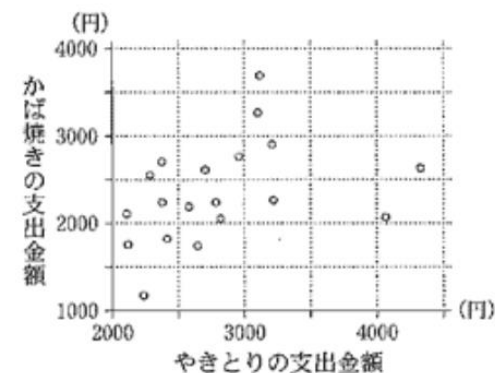


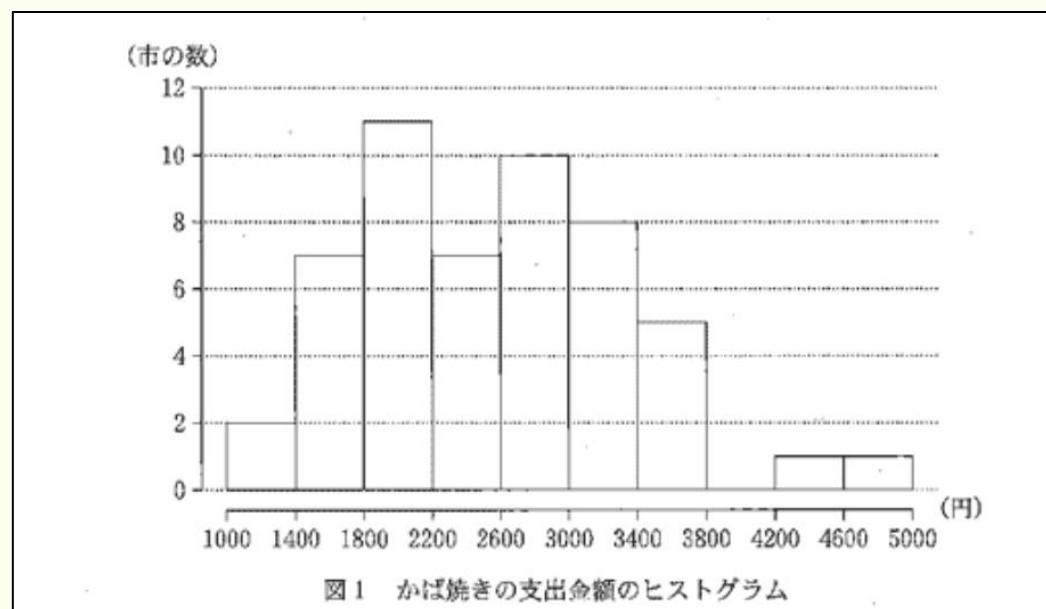
図4 地域Eにおける、やきとりとかば焼きの支出金額の散布図

表1 地域Eにおける、やきとりとかば焼きの支出金額の平均値、分散、標準偏差および共分散

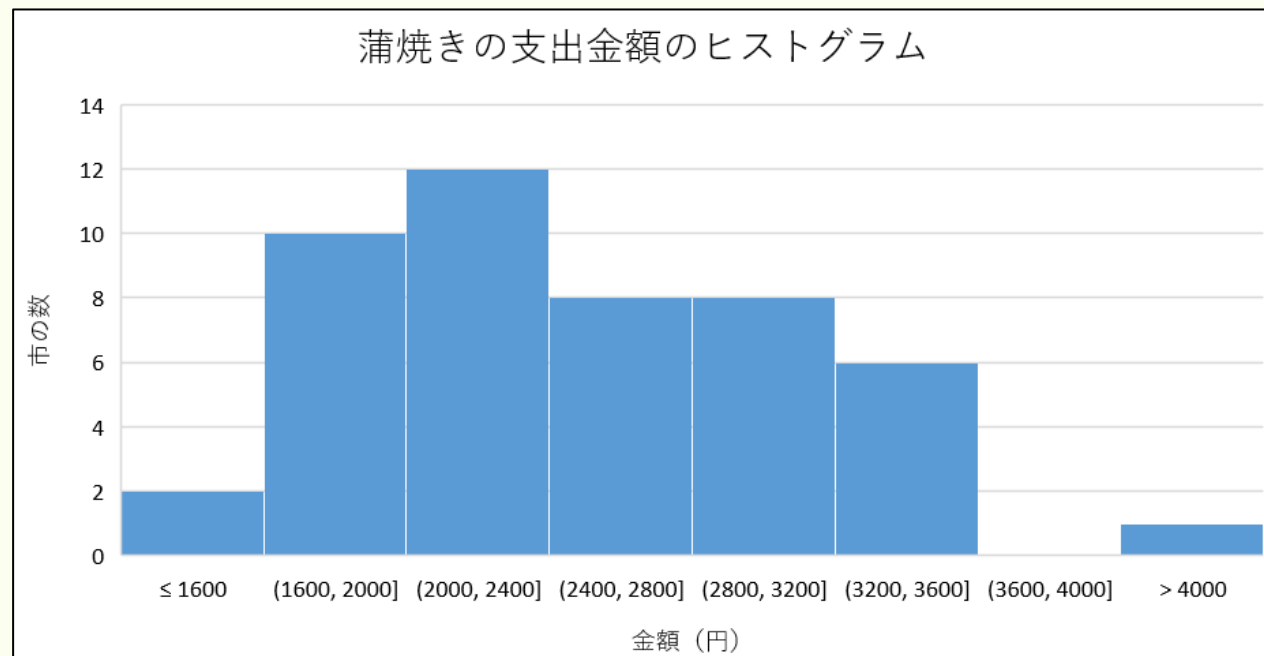
	平均値	分散	標準偏差	共分散
やきとりの支出金額	2810	348100	590	124000
かば焼きの支出金額	2350	324900	570	

ヒストグラムを作成しよう

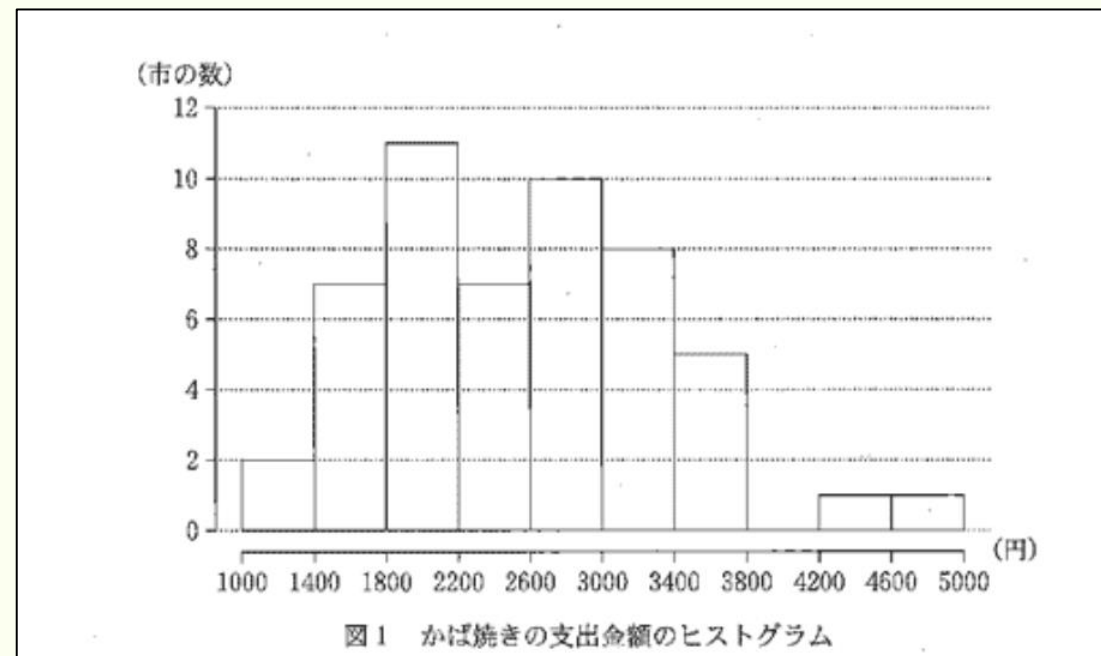
太郎さんは調理食品として、最初にうなぎのかば焼き(以下、かば焼き)に着目し、図1のように52市におけるかば焼きの支出金額のヒストグラムを作成した。ただし、ヒストグラムの各階級の区間は、左側の数値を含み、右側の数



完成！



共通テスト(再掲)



ヒストグラムが異なる理由 (SSDSEは解説・共通テストは問題文中に記載あり)

- ✓ SSDSEの元データ …… 2021-2023の平均 / 47市(県庁所在地)
- ✓ 共通テストの元データ …… 2022年 / 52市(政令指定都市等)

箱ひげ図を作成しよう

(2) 太郎さんは、東西での地域による食文化の違いを調べるために、52市を東側の地域E(19市)と西側の地域W(33市)の二つに分けて考えることにした。

(i) 地域Eと地域Wについて、かば焼きの支出金額の箱ひげ図を、図2、図3のようにそれぞれ作成した。

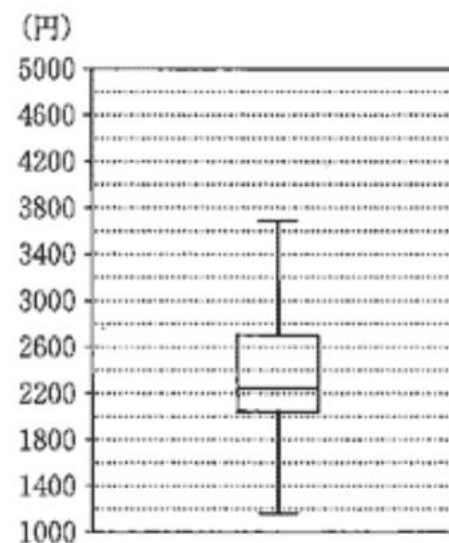


図2 地域Eにおけるかば焼きの支出金額の箱ひげ図

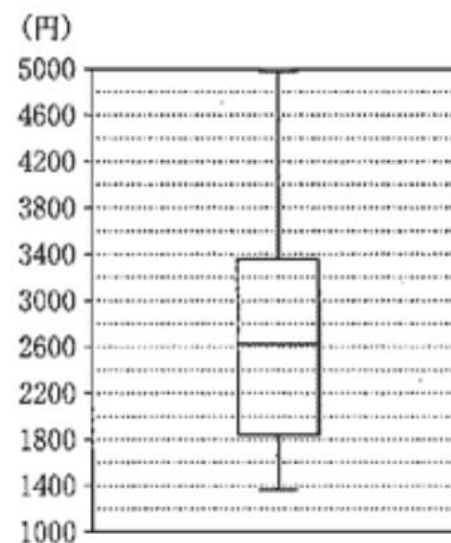
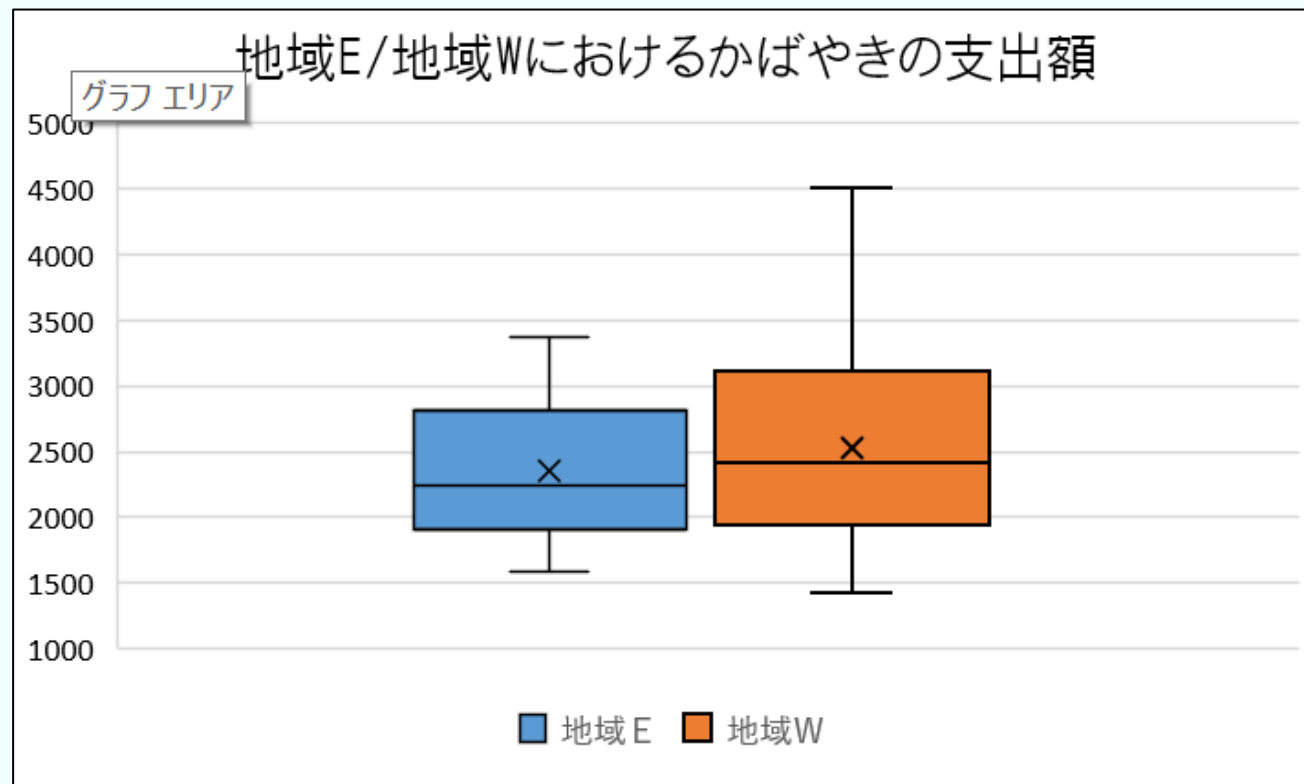


図3 地域Wにおけるかば焼きの支出金額の箱ひげ図

完成！



共通テスト(再掲)

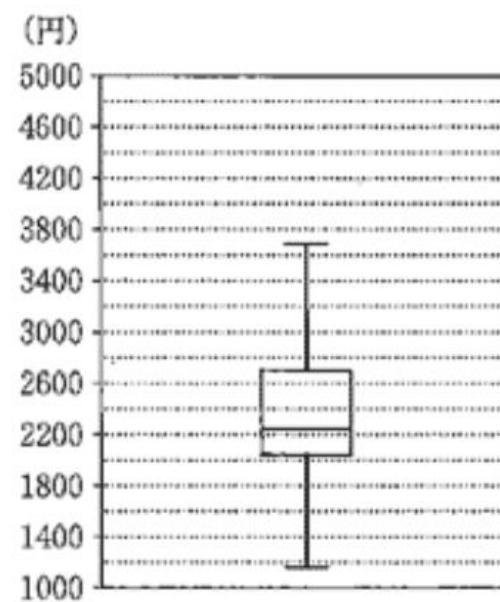


図2 地域Eにおけるかば焼きの支出金額の箱ひげ図

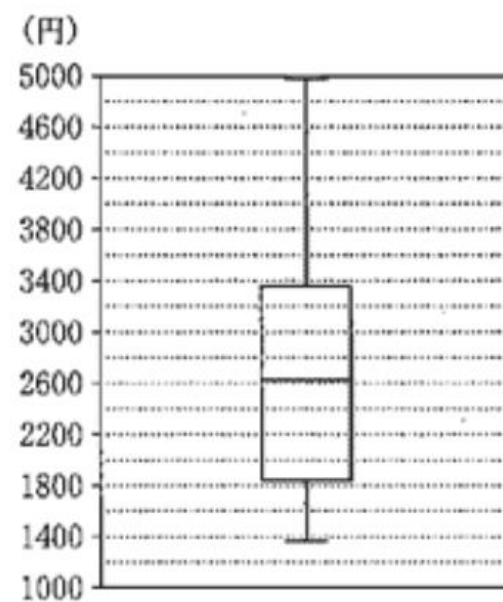


図3 地域Wにおけるかば焼きの支出金額の箱ひげ図

※ Eは18市、Wは29市に分けている

散布図を作成しよう

(3) 太郎さんは、(2)で考えた地域Eにおける、やきとりの支出金額についても調べることにした。

ここでは地域Eにおいて、やきとりの支出金額が増加すれば、かば焼きの支出金額も増加する傾向があるのではないかと考え、まず図4のように、
地域Eにおける、やきとりとかば焼きの支出金額の散布図を作成した。そ

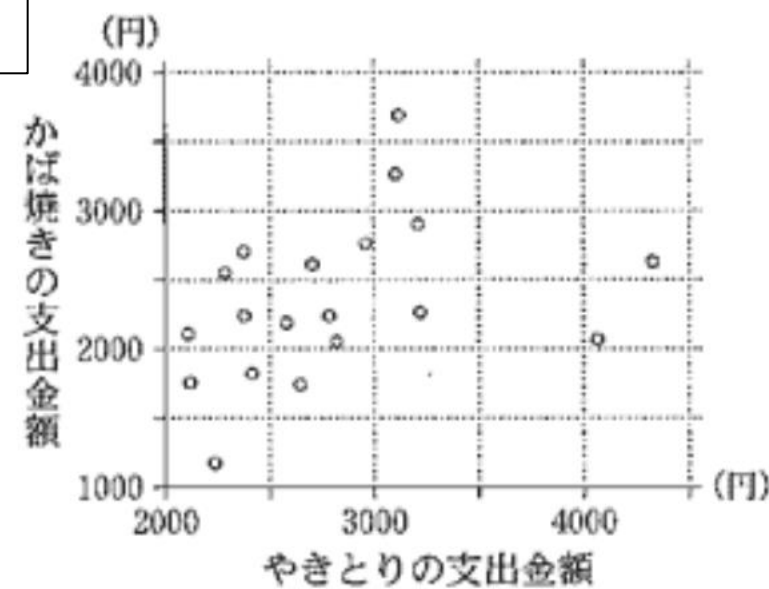
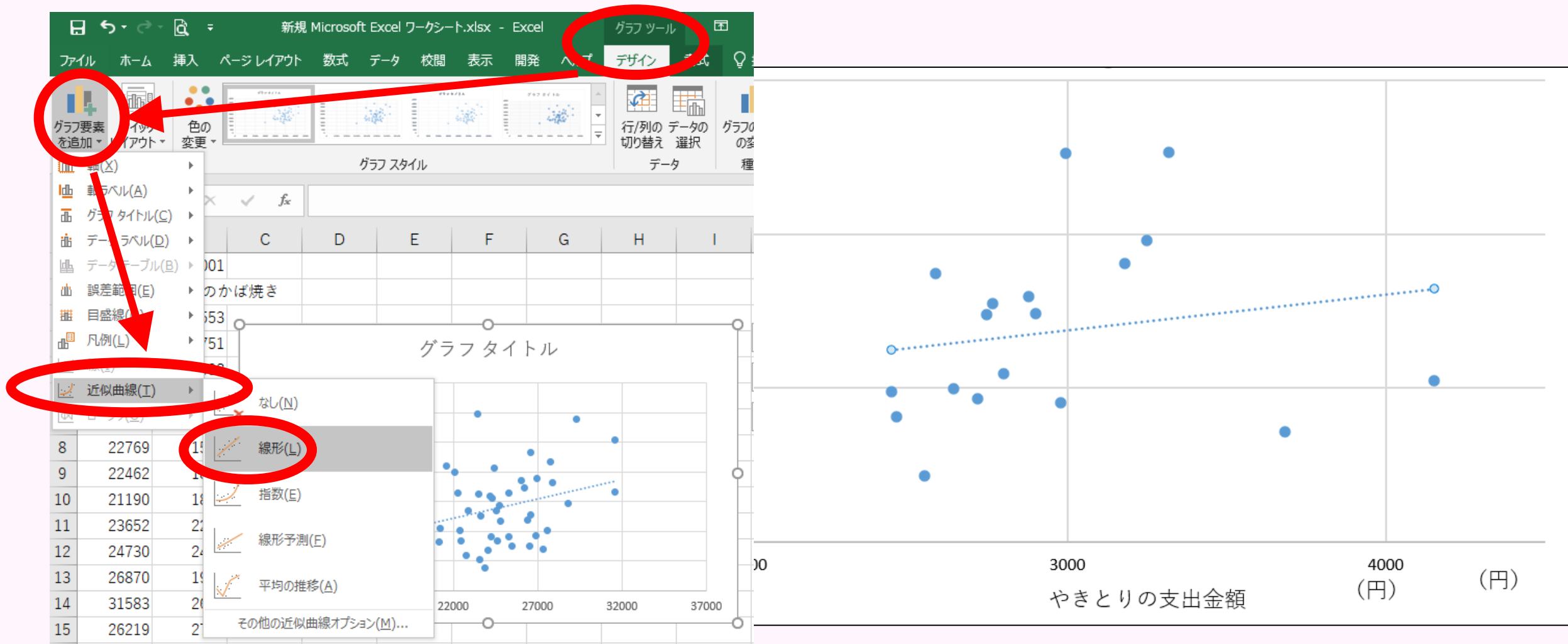


図4 地域Eにおける、やきとりとかば焼きの支出金額の散布図

4

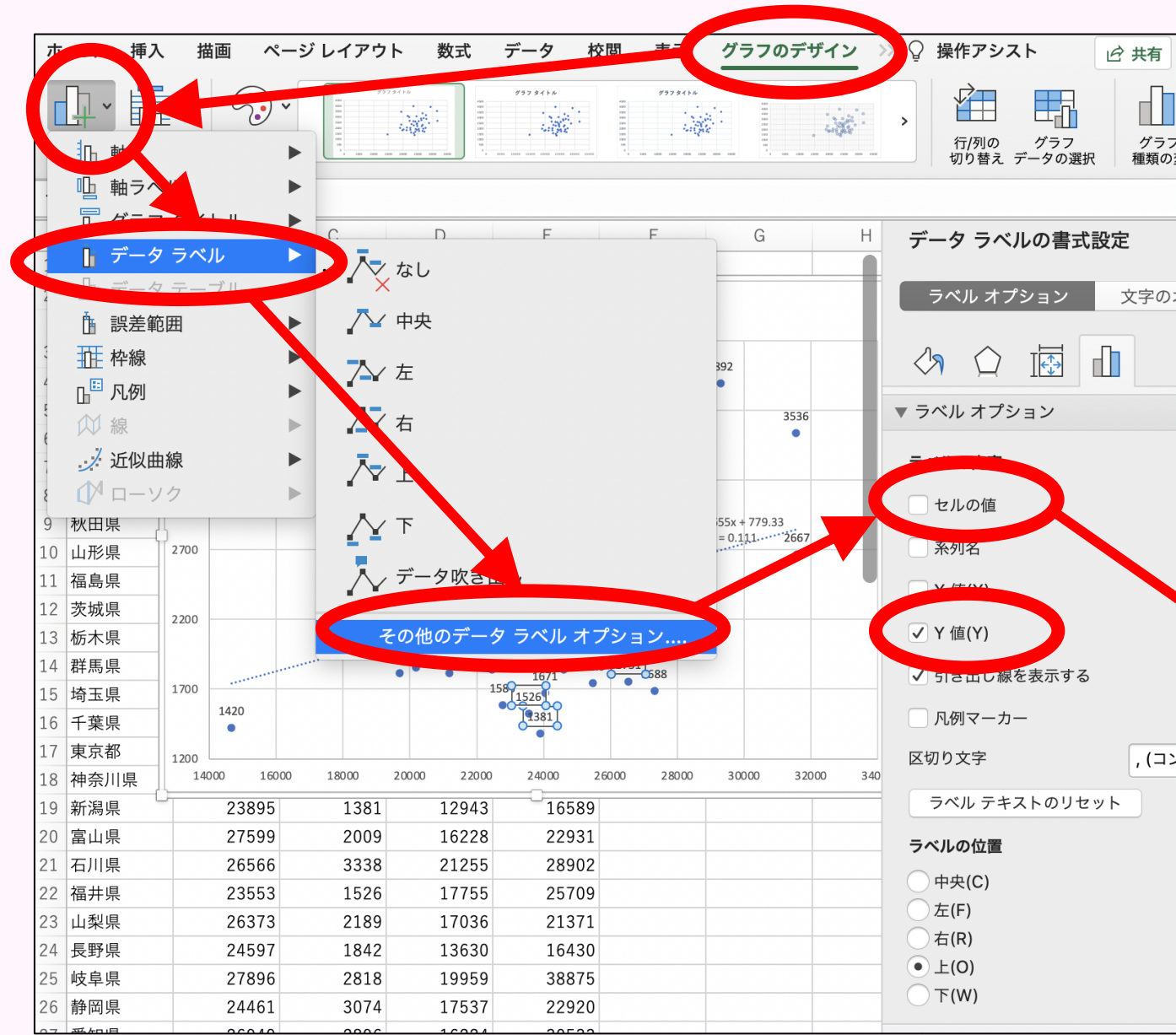
散布図：近似曲線（単回帰直線）を追加



グラフツール「デザイン」→ グラフ要素を追加 → 近似曲線 → 線形

6

散布図：県名のラベルを付ける



グラフ選択後

→ グラフのデザイン

→ グラフ要素を追加

→ データラベル

→ その他のデータラベル・・・

Y値の✓を外し、セルの値に✓

データ ラベル範囲

データ ラベル範囲の選択

キャンセル

OK

県名の入っている範囲を選択

②

基本統計量を算出しよう

後のページにまとめてある関数の表記を参考に、下の表を作成しよう

※ 共通テストとは元データが異なるので数値も異なる

		平均値	分散	標準偏差	共分散	相関係数
	やきとりの支出金額	2947	194212	441	45284.8	0.17335
	かば焼きの支出金額	2365	351398	593		

④ 夏課題レポート作成



ここに来て、ようやく、データの活用と呼べる段階へ！

- 各自、SSDSEのデータを使ってデータ分析レポートを作成
- データ分析コンペティションに応募する人は、論文を作成するよう指示
- 重回帰分析についても触れる

統計データ分析コンペティション

R4

統計活用奨励賞:都市部と地方の教育格差の要因と課題

R5

審査員奨励賞:日本の食料自給率を上げるために

審査員奨励賞:日本人の英語能力の実態とその背景

審査員奨励賞:熱中症を防ごう！

R6

統計数理賞 :子供の体力・運動能力

審査員奨励賞:鳥獣被害の原因と対策提案

今後に向けて

現状

データの可視化の練習 → 各自のテーマでデータ分析

今後

可視化されたデータの読み取りについても学ばせたい

今後に向けて

現状

化学で行った実験データを使って「t検定」を行う授業を
1コマだけ行っている

今後

さまざまな検定についても扱いたい

R6 統計数理賞

受賞生徒の論文紹介

『子供の体力・運動能力』



論文の構成

データ分析の背景と目的

データ分析の手法

結果の解釈

結論・今後に向けて

データ分析の 背景と目的

背景

全国体力・運動能力調査結果

子供たちの身体的健康を評価するための重要な指標である全国体力・運動能力調査が、**子供たちの体力が低下**している傾向を示している。

近年の環境やライフスタイルの変化

子供たちの運動能力に大きな影響を与えている。

コロナ禍の影響

運動する機会の大幅な減少により、子供たちの体力や運動能力に悪影響を及ぼしている。

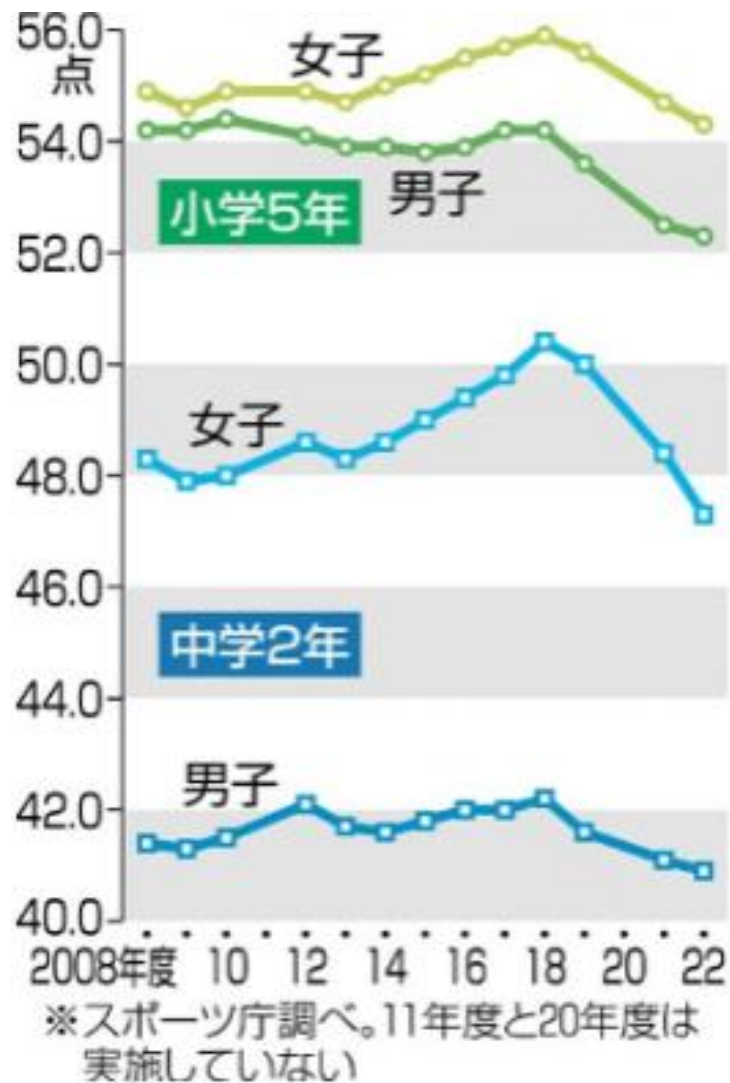


図1 全国体力テスト合計点（平均）の推移

愛知県の現状と体力づくり推進事業

愛知県は全国で最低レベル



体力向上推進プログラム



本研究を通じて

愛知県内の子供たちの

運動能力向上に寄与したい

	2021年	2022年	2023年
小5女子	47位	47位	43位
小5男子	47位	46位	47位
中2女子	45位	46位	45位
中2男子	47位	47位	46位

表1 愛知県の体力合計点都道府県別順位

データ分析の 手法

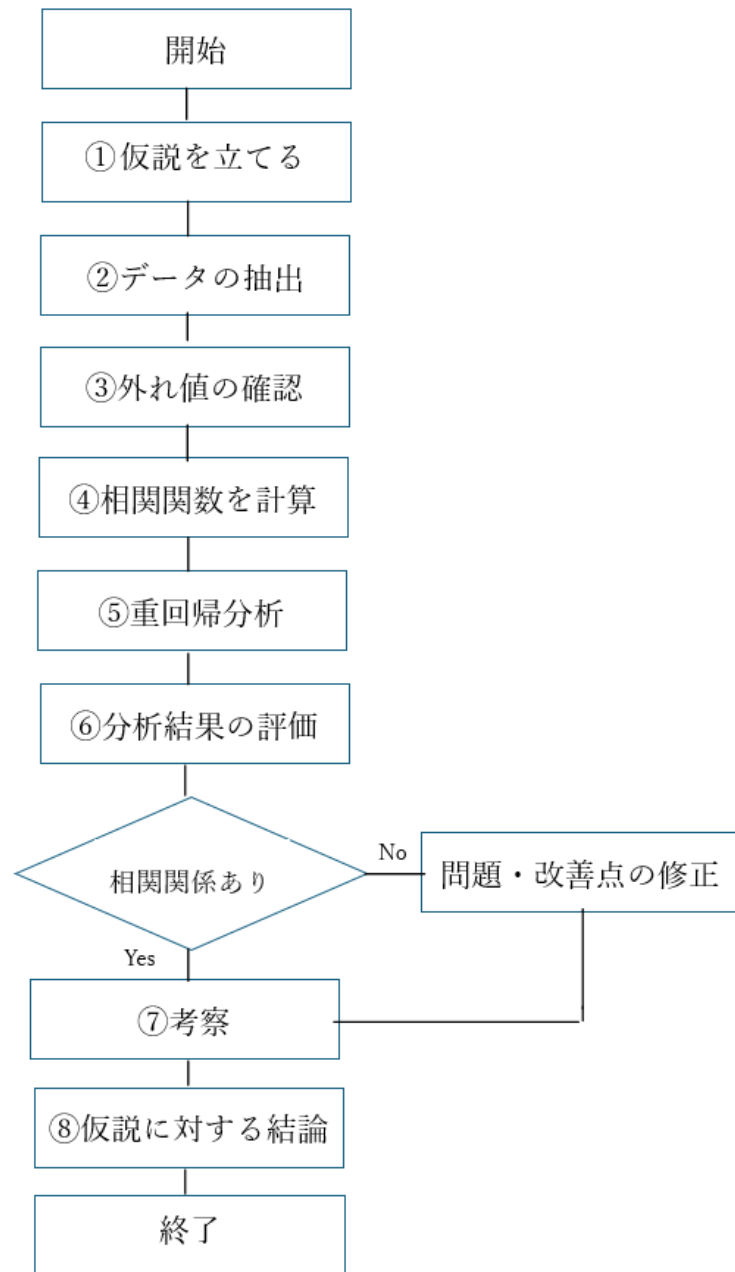


図2 研究の手順

仮説設定と要因の分類

仮説の設定

運動能力に影響を与えと考えられる要因を「運動機会」「生活習慣」「学校」「身体的要因」「環境」の5つに分け、それぞれに仮説を設定した。

要因の分類

目的変数を全国体力・運動能力調査の平均合計点とし、影響を与えていると思われるデータを説明変数として重回帰分析を行うことで、運動能力に影響を与える要因を正確に分類・考察を行った。

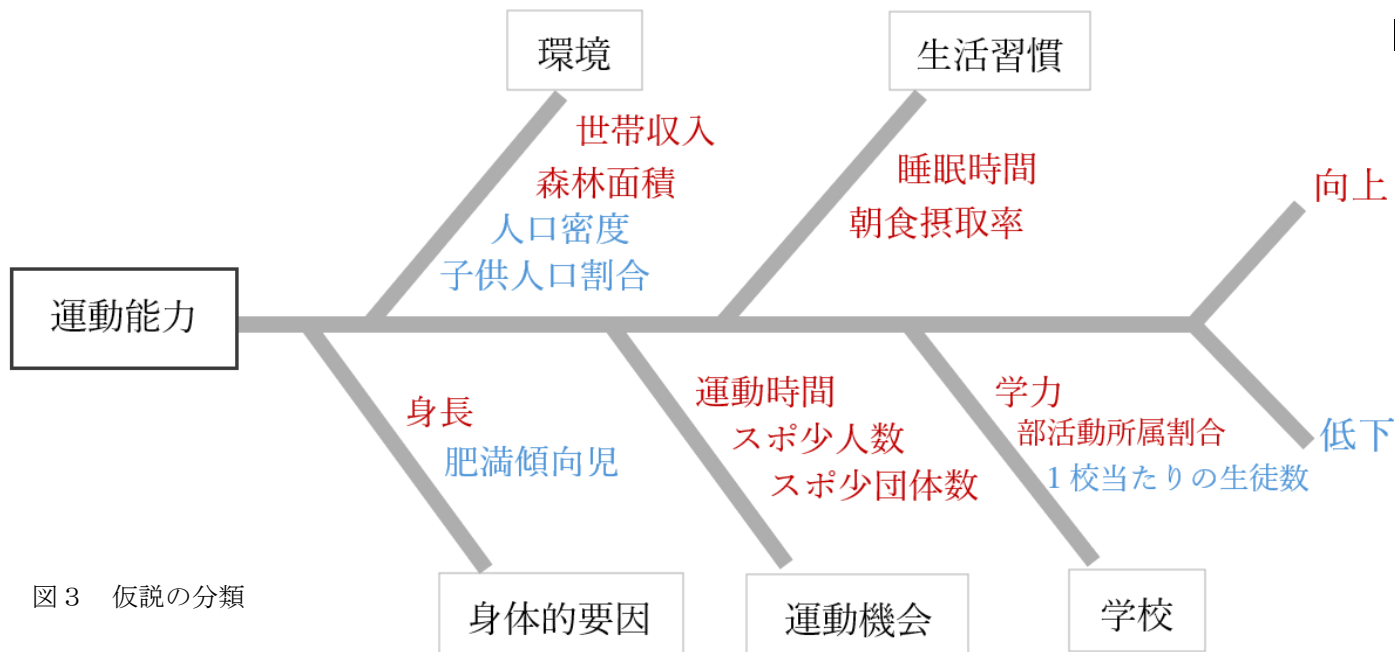


図3 仮説の分類

データの分布と外れ値の検討

都道府県ごとの特色も体力・運動能力の差につながっていると考え、外れ値となる県を除外しないこととした。

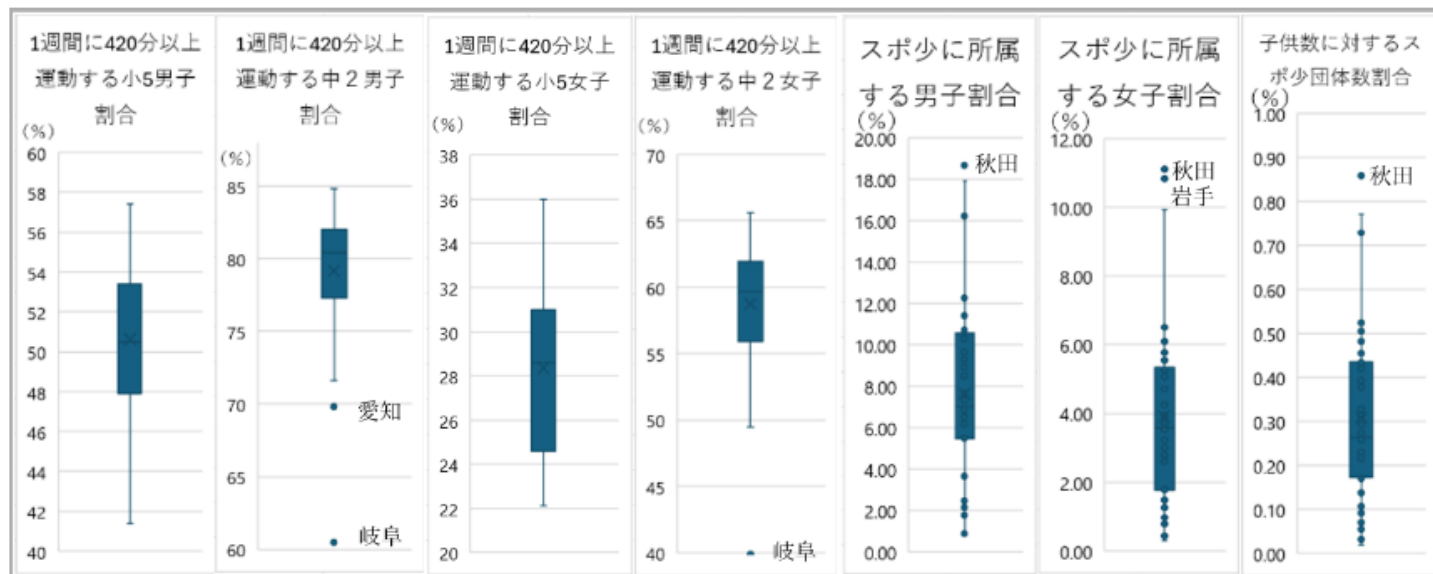
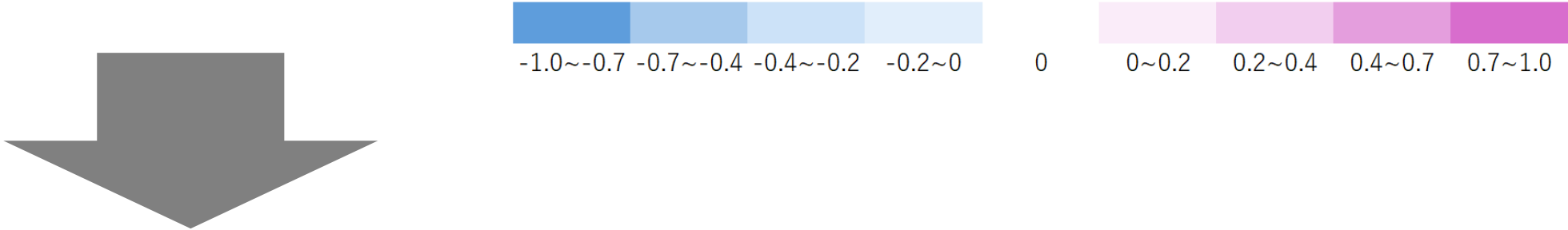


図3 データの分布と外れ値の確認 (抜粋)

相関の強い項目について重回帰分析

	運動機会			生活習慣		学校			身体的要因		環境			
	運動時間	スポ少所属	スポ少団体数	睡眠	朝食欠食率	学力	運動部	1校当たり生徒数	平均身長	肥満傾向児	県民所得	森林面積	人口密度	子供割合
小5女子	0.657	0.485	0.445		-0.111	0.383		-0.417	0.267	0.255	-0.097	0.199	-0.321	-0.285
小5男子	0.232	0.284	0.263		-0.087	0.435		-0.315	0.204	0.073	-0.116	0.192	-0.246	-0.110
中2女子	0.369	0.356	0.332	-0.220	-0.115	0.118	0.509	-0.243	0.081	0.103	-0.081	0.139	-0.324	-0.168
中2男子	0.376	0.452	0.403	-0.157	-0.065	-0.002	0.392	-0.320	0.265	0.110	-0.193	0.176	-0.372	-0.111



	切片	運動機会			生活習慣		学校			身体的要因		環境			
		運動時間	スポ少所属	スポ少団体数	睡眠	朝食欠食率	学力	運動部	1校当たり生徒数	平均身長	肥満傾向児	県民所得	森林面積	人口密度	子供割合
小5女子	-25.779	-0.035	0.659	-6.992			0.401		-0.005	0.388	0.052			-0.016	0.088
小5男子	-37.952	0.048	0.283	-6.698			0.322		-0.003	0.491				-0.023	
中2女子	42.697	0.031	0.347	-3.71	-0.054			0.066	-0.001					-0.003	
中2男子	-36.405	0.102	0.43	-8.208				-0.023	-0.002	0.455				-0.013	

結果の解釈と考察

運動習慣

運動時間 ... 正

スポ少所属人数 ... 正

スポ少団体数 ... 正

生活習慣

睡眠

・・・ 負

朝食接種率

・・・ 相関なし

学校

学力 … 正

運動部所属 … 正

1校当たり生徒数 … 負

身体的要因

平均身長 … 正

肥満児傾向 … 負 (小学生女子)

環境

県民所得 ... 相関なし

森林面積 ... 正

人口密度 ... 負

子供人口割合 ... 負

愛知県の運動能力の低さについて

<要因>

- 睡眠時間の少なさ
- スポーツ少年団の団体数や所属する割合の少なさ

<提案>

- 部活動の継続
- スポーツ少年団の充実
- 大人のボランティアの協力
- 親子で参加するスポーツイベント

最も影響を与えるのは運動習慣

- 運動系の習い事
 - 外遊びしやすい環境
- 大人の協力も必要
- 生活習慣を整える
 - 学校での勉強にも力をそそぐ

*Thank you
for your time and attention.*

参考文献

『小中学生の体力、男女とも過去最低22年度調査、コロナ禍で運動機会減る』

西日本新聞me : [nishinippon.co.jp](https://www.nishinippon.co.jp)
