

学びを活かす統計教育への取組み-SSDSEと統計データ分析コンペティション-

独立行政法人統計センター 山下雅代

総務省 統計局・総務省 統計研究研修所・(独) 統計センターでは、統計作成の専門機関として協働し、統計リテラシー向上のための支援を行っています

1. データセットの提供
SSDSE(教育用標準データセット)

SSDSE (Standardized Statistical Data Set for Education: 教育用標準データセット)

- データサイエンス教育のための汎用素材として作成・公開しているデータで、公的統計の主要な111項目のデータを全国1741市区町村別に一覧できる表形式のデータセット
- 欠測のない完備なデータセットの形 (エクセル及びCSV) で提供→すぐに分析可能

データレイアウト

縦: 1741市区町村データ × 横: 111項目

縦: 1743行 = 項目情報 (2) + 市区町村 (1741)

横: 114列 = 地域情報 (3) + データ項目 (111)

データ項目の項目名・項目情報
データ項目のコード

コード	地域情報	項目情報	A11001	A11002	A11003	A11004	...	S2003	S2004
A01100	北海道	旭川市	1952256	910414	1041762	...	226	25037	
A01200	北海道	札幌市	2626716	1201761	1456651	...	45	3094	
A01300	北海道	小樽市	1272424	54985	66939	...	21	1294	
A02700	青森県	青森市	1184	630	555	...	1	35	
A02800	青森県	八戸市	3996	2057	1941	...	0	0	
A02900	青森県	三好市	1943	1147	986	...	1	35	

市区町村コード 都道府県名 市区町村名
地域情報 データ項目 (111)

横: 111項目一覧

解きたい課題に関連する変数を選んで使う

【人口・世帯】	人口総数 (性別、年、5)	【経済指標】	中学校就学率
日本人人口 (性別、年、5)	外国人人口	産業別就労率 (産業大分類(17))	中学校卒業後
15歳未満人口 (性別、年、5)	出生数	第1次産業従事者数	高等学校
15-64歳人口 (性別、年、5)	死亡数	第2次産業従事者数	高等学校卒業
75歳以上人口 (性別、年、5)	転入人口	第3次産業従事者数	高等学校卒業
転出人口	外国人人口	第1次産業就労率	【文化・スポーツ】
【労働指標】	雇用者人口	第2次産業就労率	国民文化祭
就業率	パート・アルバイト人口	第3次産業就労率	文化フェスティバル
1次産業就労率	正規雇用者人口	【教育】	幼稚園
2次産業就労率	パート・アルバイト人口	保育所	小学校教育
3次産業就労率	正規雇用者人口	小学校	小学校教育
【行政指標】	町長選挙結果	中学校	中学校卒業後
町長選挙結果	町長選挙結果	高等学校	高等学校卒業
町長選挙結果	町長選挙結果	高等学校卒業	高等学校卒業
町長選挙結果	町長選挙結果	高等学校卒業	高等学校卒業
町長選挙結果	町長選挙結果	高等学校卒業	高等学校卒業

SSDSE2019版

SSDSE2018版を最新のデータに更新するだけでなく、都道府県別時系列のデータセットを加え、2019年6月末公表予定

2. Good Practiceの共有
統計データ分析コンペティション

★趣旨
教育用標準データセット (SSDSE) を用いた統計分析の論文を募集し、アイデアと解析力を競うことで、高校生、大学生等の統計リテラシーの向上を図ります。

★エントリー期間
2019年 (令和元年) 7月1日 (月) から8月9日 (金) まで

★論文締切
[高校生部] 2019年9月17日 (火) 18時まで
[大学生・一般部] 2019年9月 9日 (月) 18時まで

★賞及び副賞等
表彰優秀作品には、総務大臣賞、優秀賞、統計数理賞、統計活用奨励賞と各副賞が授与されます。また、受賞論文が統計専門誌等に掲載されます。



コンペティションの受賞による効果

広島大学附属中・高等学校 数学科 橋本三朗先生講演資料(橋本, 2019)よりまとめ
2018年 高校生の部 総務大臣賞受賞「本当に日本の医療は危機的状況にあるのか?」
広島大学附属高等学校 大段 利々子さん

- ★ 専門家の発展的コメント、他の受賞作品との出会い、「統計」誌への投稿 → 探索的な学びのサイクルが回る
- ★ 次回も挑戦したい → チャレンジ精神・積極性が育まれる
- ★ 自分で決めたことは最後までやり遂げられる → 自信・自己肯定感が高まる
- ★ 自分もやってみよう → 周りの生徒のやる気を引き出す

SSDSE2018年, 2019年版どちらを使っても分析でも応募可

3. 探究型統計教材の提供 SSDSEの学びの事例群

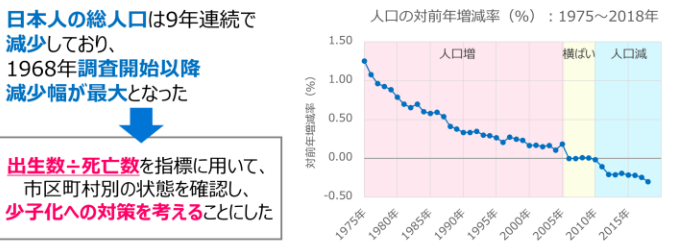
教材の開発方針
「少子化」問題をテーマに基本的概念 (知識) と活用方法の習得を志向

開発済教材の一覧: 月刊「統計」誌 2019年4月号~2020年3月号連載*

学習内容	教材名	概要	想定学年
① 四分位数・箱ひげ図	散らばり方を比較しよう	四分位数の導出と、箱ひげ図による複数の分布の比較	現: 数1 新: 中2
② 箱ひげ図・ヒストグラム・層別	外れ値に着目しよう	外れ値に着目した問題発見、及び層別と指数化による特徴の抽出	中2~
③ 相関係数・散布図	相関分析をしてみよう	疑似相関の確認と、外れ値の処理	数1
④ 層別	必要な汗はかこう	インターネット調査による層別と外れ値市区町村の分類	中1~

* (一財) 日本統計協会発行、SSDSE及び統計データ分析コンペティションのwebサイトでも紹介予定

問題場面: 少子化



ヒストグラムの4つの着眼点と問題解決の流れ

① 分布の中心の把握

日本全国1737市区町村の出生数÷死亡数のヒストグラム

中央値 (死者10名に対し出生者5名)

散らばり方が左右非対称なとき

平均値 Or 中央値

分布に歪みがある ⇒ 中央値 (あるいは最頻値) を用いて説明

② 割合の把握

日本全国1737市区町村の出生数÷死亡数のヒストグラム

出生数 < 死亡数 → 人口減少
出生数 > 死亡数 → 人口増加

人口が減少している市区町村の割合が高い ⇒ 問題の大きさの確認

③ 外れ値の把握

日本全国1737市区町村の出生数÷死亡数のヒストグラム

出生数 >> 死亡数 → 人口増加が顕著な市区町村

外れ値の特徴を見出せば、仮説が得られる ⇒ 問題解決のための糸口

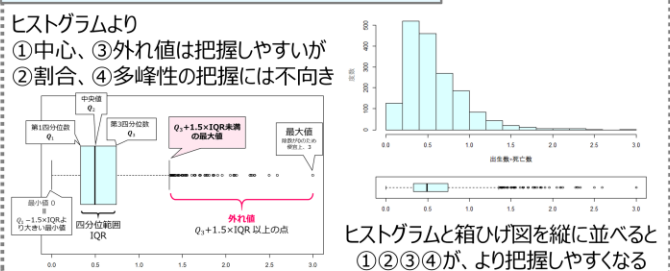
④ 多峰性の有無の把握

石川県の出生数÷死亡数のヒストグラム

層1, 層2, 層3

野々市市、川北町はどんな市町か? ⇒ 3つに層別 (分類)

補足: 外れ値を示した箱ひげ図



地図グラフで確認

金沢市と小松市の間にある面積の小さい市町が外れ値

金沢市: 県庁所在地
小松市: 小松製作所の本拠地
日本有数の企業城下町

仮説: 野々市市と川北町の特徴を調べ、少子化対策のヒントがある

政府統計ポータルサイト e-Stat の活用のススメ

- 仮説検証に必要なデータが揃う
- 統計GIS JSTAT Map を使えば地図グラフも簡単に描ける

引用・参考文献
橋本三朗(2019). SSDSE統計データ分析コンペティションに参加して—論文作成による高校1年生の思考の深まり—. 統計教育実践研究, 特別号, 71-72.
教育用標準データセット(SSDSE) https://www.nstac.go.jp/ssdse/ (2019.5.16最終確認)
統計データ分析コンペティション https://www.nstac.go.jp/statcomp/ (2019.5.16最終確認)
山下雅代・橋本三朗・飯島忠也(2019). 教育用標準データセット(SSDSE)による探究型統計教育の促進—総務省統計コミュニティの試み—. 日本数学会誌, 101(3), 39-46.