



独立行政法人  
統計センター

# SSDSEと統計データ分析コンペティション の概要

信頼にこたえてつくる確かな統計

2023年3月12日

田中雅行

独立行政法人統計センター

第20回 統計教育の方法論ワークショップ

セッションIII: 統計データ分析コンペティション  
を用いた統計教育

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

- ・ データ分析のための汎用素材として、統計センターが提供しているデータセット
- ・ 政府統計の地域別データを手軽に利用できるよう編集して提供
- ・ 縦に地域、横にデータ項目を並べた、2次元の表形式データ

### ○ 2018年に提供を開始し、現在は5種類のSSDSEを無償提供

- ・ SSDSE-市区町村 (SSDSE-A) 市区町村別、多分野データ
- ・ SSDSE-県別推移 (SSDSE-B) 都道府県別、時系列、多分野データ
- ・ SSDSE-家計消費 (SSDSE-C) 都道府県庁所在市別、家計消費データ
- ・ SSDSE-社会生活 (SSDSE-D) 都道府県別、自由時間活動・生活時間データ
- ・ SSDSE-基本素材 (SSDSE-E) 様々な分野の都道府県別データ

⇒ 毎年、最新データに更新

⇒ 今後も内容を充実 (新たな SSDSE の開発・提供など)

### ○ SSDSE を用いた統計データ分析コンペティションを毎年度開催

## SSDSEの4つの特徴

### 簡便性

- ・ 無償で容易に使うことができる
  - ⇒ 一括ダウンロードし必要なデータを取捨選択
- ・ 特別な前処理が不要
  - ⇒ 全セルに数値が入っている
  - ⇒ **欠測値がない「完備データ」**

### 親近性

- ・ 自分に関係する地域のデータが含まれる
  - ⇒ 誰でもなじみがある「**自分たちのデータ**」

### 具体性

- ・ 個々のデータの意味が分かり議論できる
  - ⇒ **政府統計から抽出・編集した「実データ」**

### 多様性

- ・ 幅広い分野のデータが含まれている
  - ⇒ 様々な課題抽出、多様な分析が可能
  - ⇒ 自由度の高い「**標準データ**」

# 1. SSDSEの概要

## SSDSE

- ・ EXCEL と CSV で 提供



統計を活かす

公的統計の二次的利用サービス

- 調査票情報のオンサイト利用
- オナーメード集計の利用

利用

ために

クション

データセット)

標準データセット)

### SSDSE (教育用標準データセット)



- 2023年2月27日 **NEW**  
SSDSE-社会生活(SSDSE-D)の改訂版を公開しました。

SSDSE(教育用標準データセット: Standardized Statistical Data Set for Education)は、データ分析のための汎用素材として、独立行政法人統計センターが作成・公開している統計データです。主要な公的統計を地域別に一覧できる表形式のデータセットで、直ちにデータ分析に利用することができます。データサイエンス演習、統計教育などにご活用ください。

SSDSEを用いた統計データ分析コンペティションを、高校生、大学生等を対象に開催しています。

### SSDSEの一覧

名称	最新版ファイル名	内容
----	----------	----

### SSDSE-市区町村(SSDSE-A)

- 様々な分野の市区町村別データを集めたデータセットです。(1741市区町村×多分野124項目)  
(出典)総務省統計局「統計でみる市区町村のすがた(社会・人口統計体系)2022」

	SSDSE (EXCEL)	SSDSE (CSV)	SSDSEの解説
2022年版 2022年6月30日公開	<a href="#">SSDSE-A-2022</a> (EXCEL:1,274KB) 	<a href="#">SSDSE-A-2022</a> (CSV:914KB) 	<a href="#">SSDSE-市区町村の解説</a> (PDF:1,015KB)  

### SSDSE-県別推移(SSDSE-B)

- 様々な分野の都道府県別・時系列データを集めたデータセットです。(47都道府県×12年次×多分野107項目)  
(出典)総務省統計局「統計でみる都道府県のすがた(社会・人口統計体系)2022」

	SSDSE (EXCEL)	SSDSE (CSV)	SSDSEの解説
2022年版 2022年4月28日公開、2023年2月9日修正	<a href="#">SSDSE-B-2022</a> (EXCEL:412KB) 	<a href="#">SSDSE-B-2022</a> (CSV:344KB) 	<a href="#">SSDSE-県別推移の解説</a> (PDF:882KB)  

2023年2月9日の修正は、一部欠落のあった項目名を補足したもので、数値データ及び「解説」の修正はありません。

## SSDSE の解説



- ・ 利用する際の留意点などを記載

<https://www.nstac.go.jp/use/literacy/ssdse/>

# 1. SSDSEの概要

## (例) SSDSE-県別推移の解説



- データの出典、単位、年次や、利用する際の留意点など、分析する上で必要な情報を掲載

### SSDSE-県別推移の解説

SSDSE-B-2022



- SSDSE (教育用標準データセット: Standardized Statistical Data Set for Education) データサイエンス教育のための汎用素材として公開している統計データです。
- 主要な公的統計の地域別データを表形式に編集したもので、欠測データがないので、ダウンロード直ちにデータ分析に使用できます。
- SSDSE-県別推移 (SSDSE-B) は、様々な分野の都道府県別・時系列データを集めたデータセットです。  
(出典) 総務省統計局「統計でみる都道府県・市区町村のすがた (社会・人口統計体系)」

#### SSDSEを授業や講義でお使いになる先生方へ

- SSDSEを配布する際には、この「解説」も一緒に配布していただき、データの出典や単位についても意識してお使いいただけますよう、お願いいたします。

### データのレイアウト

縦 (行の数) 項目情報 (2) + 都道府県 (47) × 年次 (12) = 566

横 (列の数) 年次情報 (1) + 地域情報 (2) + データ項目 (107) =

※ 前年に比べ、データ項目が2項目増加しました (3ヶ)

項目情報	SSDSEのID情報 (SSDSE-B-2022)		地域情報		データ項目 (107)				L322109	
	項目コード	項目名	SSDSE-B-2022 Code	Prefecture	A1101	A110101	A110102	A1102		
12年次	年度	地域コード	都道府県	総人口	総人口 (男)	総人口 (女)	日本人人口			
				2019 R01000	北海道	5250000	2472000	2778000	5211000	293
				2018 R01000	北海道	5286000	2489000	2797000	5253000	269

### 別表 SSDSE-県別推移のデータ一覧

- ◎項目コードをクリックすると「社会・人口統計体系」の項目定義のサイトが表示されます。
- 本SSDSEに収録されていない関連データの入手には、出典の名称を参考に検索してください。

項目コード	項目名	単位	出典	e-Statで入手できる年次の範囲
<b>A</b>	<b>人口・世帯</b>			
A1101	総人口	人	・国勢調査実施年 (5年毎: 西暦0又は5)	1975~2019
A110101	総人口 (男)	人	総務省統計局	
A110102	総人口 (女)	人	「国勢調査」人口等基本集計	
A1102	日本人人口	人	・国勢調査実施年以外	
A110201	日本人人口 (男)	人	総務省統計局	
A110202	日本人人口 (女)	人	「人口推計」	
A1301	15歳未満人口	人		
A130101	15歳未満人口 (男)	人		
A130102	15歳未満人口 (女)	人		
A1302	15~64歳人口	人		
A130201	15~64歳人口 (男)	人		
A130202	15~64歳人口 (女)	人		
A1303	65歳以上人口	人		
A130301	65歳以上人口 (男)	人		
A130302	65歳以上人口 (女)	人		
A4101	出生数	人	厚生労働省政策統括官 (統計・情報政策担当)	1980~2020
A410101	出生数 (男)	人	「人口動態統計」	
A410102	出生数 (女)	人		
A4103	合計特殊出生率	-		1980~2019
A4200	死亡数	人		1975~2020
A420001	死亡数 (男)	人		
A420002	死亡数 (女)	人		
A5101	転入者数 (日本人移動者)	人	総務省統計局	1975~2020
A510101	転入者数 (日本人移動者) (男)	人	「住民基本台帳人口移動報告」	
A510102	転入者数 (日本人移動者) (女)	人		
A5102	転出者数 (日本人移動者)	人		
A510201	転出者数 (日本人移動者) (男)	人		
A510202	転出者数 (日本人移動者) (女)	人		
A9101	婚姻件数	組	厚生労働省政策統括官 (統計・情報政策担当)	1975~2020
A9201	離婚件数	組	「人口動態統計」	
<b>B</b>	<b>自然環境</b>			
B4101	年平均気温	℃	気象庁観測部	1975~2020
B4102	最高気温 (日最高気温の月平均の最高値)	℃	「過去の気象データ」	

# 1. SSDSEの概要

## SSDSE-県別推移：都道府県別、時系列、多分野データ



SSDSE-B-2	Code	Prefecture	A1101	A110101	A110102	A1102	A110201	A110202	A1301	A130101	A130102
年度	地域コード	都道府県	総人口	総人口 (男)	総人口 (女)	日本人人口	日本人人口 (男)	日本人人口 (女)	15歳未満人 口	15歳未満人 口(男)	15歳未満人 口(女)
2019	R01000	北海道	5250000	2472000	2778000	5211000	2455000	2756000	565000	289000	276000
2018	R01000	北海道	5286000	2489000	2797000	5253000	2475000	2778000	577000	295000	282000
2017	R01000	北海道	5320000	2506000	2814000	5292000	2494000	2797000	588000	301000	288000
2016	R01000	北海道	5352000	2521000	2830000	5327000	2511000	2816000	600000	306000	294000
2015	R01000	北海道	5381733	2537089	2844644	5348768	2522980	2825788	608296	310387	297909
2014	R01000	北海道	5410000	2551000	2859000	5390000	2543000	2847000	621000	316000	304000
2013	R01000	北海道	5438000	2565000	2873000	5419000	2558000	2861000	630000	321000	309000
2012	R01000	北海道	5465000	2580000	2886000	5446000	2572000	2874000	640000	326000	314000
2011	R01000	北海道	5488000	2593000	2896000	5470000	2585000	2884000	650000	331000	318000
2010	R01000	北海道	5506419	2603345	2903074	5482650	2593193	2889457	657312	335353	321959
2009	R01000	北海道	5524000	2612000	2911000	5506000	2605000	2901000	663000	338000	325000
2008	R01000	北海道	5548000	2626000	2922000	5530000	2619000	2911000	675000	344000	331000
2019	R02000	青森県	1246000	585000	661000	1240000	583000	657000	133000	68000	65000
2018	R02000	青森県	1263000	593000	670000	1258000	591000	666000	137000	70000	67000
2017	R02000	青森県	1278000	600000	678000	1274000	599000	675000	141000	72000	69000
2016	R02000	青森県	1293000	608000	686000	1290000	606000	683000	145000	74000	71000
2015	R02000	青森県	1308265	614694	693571	1302132	612113	690019	148208	75661	72547
2014	R02000	青森県	1323000	622000	701000	1320000	620000	699000	155000	79000	76000
2013	R02000	青森県	1337000	628000	708000	1334000	627000	706000	159000	81000	78000
2012	R02000	青森県	1350000	635000	716000	1347000	634000	714000	164000	83000	80000

# 1. SSDSEの概要

## ◎ 現行の SSDSE の比較表

名称	SSDSE-市区町村	SSDSE-県別推移	SSDSE-家計消費	SSDSE-社会生活	SSDSE-基本素材
内容	市区町村別、 多分野データ	都道府県別、 時系列、 多分野データ	都道府県庁所在市別、 家計消費データ	都道府県別、 自由時間活動・ 生活時間データ	都道府県別、 他分野データ
表のサイズ	縦 1744 × 横 127	縦 566 × 横 110	縦 50 × 横 229	縦 146 × 横 124	縦 51 × 横 92
地域数	1741市区町村	47都道府県	全国 + 47県庁所在 市	全国 + 47都道府 県	全国 + 47都道府 県
元になる統計	総務省「統計でみる都道府県・市区町村のすがた（社会・人口統計体系）」		総務省「家計調査」	総務省「社会生活基本調査」	総務省「統計でみる都道府県・市区町村のすがた（社会・人口統計体系）」
データ項目数	124	107	226	121	90
データの年次	各項目、最新の1 年次	2008～2019年 (12年次)	2019年～2021年の 平均値	2021年	各項目、最新の 1年次
提供開始年	2018年	2019年	2020年	2021年	2022年

※ 毎年、最新データに更新

今後とも内容の充実を図る（新たな SSDSE の開発・提供など）

## 2. 統計データ分析コンペティションの概要

- 我が国の次代を担う高校生、大学生等の統計データの利活用マインドと分析技術のかん養を図るとともに、公的統計の利活用を推進するため、2018年度に初めて開催し、2022年度に5回目を実施。
- 地域別の統計をまとめたSSDSE（教育用標準データセット：Standardized Statistical Data Set for Education）を用いた統計データ分析の論文を募集し、そのアイデアと解析力を競うコンペティションであり、成績優秀者には、総務大臣賞等の表彰を行う。

## 2. 統計データ分析コンペティションの概要

**共催：総務省統計局・独立行政法人統計センター・**

**大学共同利用機関法人情報・システム研究機構統計数理研究所・**

**一般財団法人日本統計協会**

**後援：文部科学省・国立研究開発法人科学技術振興機構（JST）・**

**全国高等学校長協会・一般社団法人日本統計学会・**

**全国統計教育研究協議会**

**募集部門：高校生の部・・・高校、高専（1～3年次）の生徒**

**大学生・一般の部・・・短大、高専（4、5年次、専攻科）、**

**大学、大学院の学生、一般**

# 2. 統計データ分析コンペティションの概要

## 2022年の主な日程

- 5月10日 エントリー及び論文募集開始
- 8月10日 エントリー締切
- 9月1日 論文提出締切 大学生・一般の部
- 9月9日 論文提出締切 高校生の部
- 10月18日（統計の日）受賞論文の発表
- 12月7日（全国統計大会）受賞者表彰式



総務省統計局 独立行政法人 統計センター 大学共同利用機関法人 情報・システム研究機構 統計数理研究所 The Institute of Statistical Mathematics 一般財団法人 日本統計協会 Japan Statistical Association

# 統計データ分析 コンペティション 2022

高校生、大学生等の皆さんの**統計分析**における  
アイデアと技術を競います

## 論文募集

**エントリー期間**  
令和4年5月10日（火）  
～8月10日（水）

**論文締切**

[大学生・一般の部]  
令和4年9月 1日（木）

[高校生の部]  
令和4年9月 9日（金）

最優秀作品には  
**総務大臣賞と副賞**  
が贈られます  
このほか、受賞論文が  
専門誌等に掲載されます

第5回統計データ分析コンペティション  
詳細は、下記ウェブサイトをご覧ください。  
<https://www.nstac.go.jp/statcompe/>



**共催** 総務省統計局 ・ 独立行政法人 統計センター  
大学共同利用機関法人 情報・システム研究機構 統計数理研究所 ・ 一般財団法人 日本統計協会

**後援** 文部科学省 ・ 国立研究開発法人 科学技術振興機構 (JST)  
全国高等学校長協会 ・ 一般社団法人 日本統計学会 ・ 全国統計教育研究協議会

# 2. 統計データ分析コンペティションの概要

【別紙】

報道資料



MIC Ministry of Internal Affairs and Communications

令和4年4月15日

## 「統計データ分析コンペティション2022」を開催します！

総務省は、高校生、大学生等の統計の有用性への理解と統計データの利活用拡大のため、独立行政法人統計センター等との共催により、「統計データ分析コンペティション2022」を開催します。

### ○ コンペティションの概要

高校生、大学生等を対象に、地域別の統計をまとめた「教育用標準データセット (SSDSE)」を用いた統計データ分析の論文を募集し、そのアイデアと解析力を競うコンペティションです。平成30年度から開催しており、令和4年度は5回目となります。

(注) SSDSEは、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セットです。

### ○ 応募資格

【高校生の部】

高等学校、中等教育学校（後期課程）及び高等専門学校（1～3年次）の生徒

【大学生・一般の部】

短期大学、高等専門学校（4、5年次、専攻科）、大学及び大学院の学生並びに一般（統計分析等の学習を目的とする方）

### ○ 募集期間

エントリー期間 令和4年5月10日（火）～8月10日（水）

論文の提出期限  
9月1日（木）（大学生・一般の部）  
9月9日（金）（高校生の部）

### ○ 表彰

審査の結果、優秀な論文には表彰状及び副賞を授与します。

（別添資料）

別紙 「統計データ分析コンペティション2022」の開催

（連絡先）

【報道発表の内容について】

統計局統計情報利用推進課

担当：和田課長補佐 山田主査

電話：03-5273-1023（直通）

E-mail：y-senryaku\_atmark\_soumu.go.jp

【応募方法とSSDSE等について】

（独）統計センター技術研究開発課

担当：田中統括統計職 堀川統計専門職

電話：03-5273-1368（直通）

E-mail：statcompe\_atmark\_nstac.go.jp

※スパムメール防止のため、「@」を「\_atmark\_」と表示しています。メールを送る際に「\_atmark\_」を「@」に直して入力してください。



「統計データ分析コンペティション2022」の開催

高校生、大学生等が統計の有用性を理解し、統計データを適切に活用できるよう、統計データの分析に関するコンペティションを令和4年度も開催します。奮って御応募ください！！

### ▶ 概要

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

令和4年5月10日（火）から8月10日（水）まで

### ▶ 論文締切

【大学生・一般の部】 令和4年9月1日（木）

【高校生の部】 令和4年9月9日（金）

### ▶ 表彰

優秀な論文には、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。

### ▶ 共催

- ・総務省統計局
- ・独立行政法人 統計センター
- ・大学共同利用機関法人 情報・システム研究機構 統計数理研究所
- ・一般財団法人 日本統計協会

### ※教育用標準データセット (SSDSE) とは

SSDSE: Standardized Statistical Data Set for Education

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、様々な分野（人口・世帯、経済基盤、教育、健康・医療、福祉・社会保障など）の公的統計を、市区町村別又は都道府県別にまとめた、表形式のデータセットです。

（データセットは、A.市区町村別、B.都道府県別の地域データに加えて、C.家計消費データ、D.自由時間活動・生活時間データの4種類を提供中。順次、更新・追加予定）

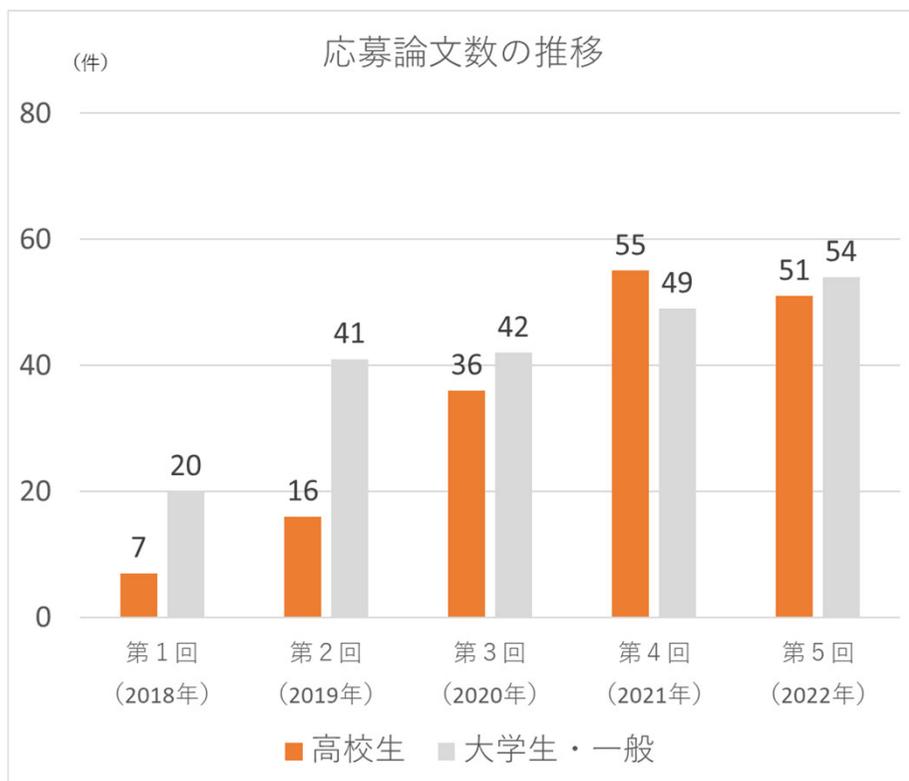
詳細については、当コンペティションのウェブサイトをご覧ください。

<https://www.nstac.go.jp/statcompe/>

# 3. 統計データ分析コンペティションの応募状況

応募論文数は増加傾向

応募者還元策としての個別講評の試行



参加者からの要望に応え、  
**高校生の全応募論文に個別講評を実施**

論文一つひとつの内容に即した講評を  
 高校生の全応募者（51件）に対して  
 作成し、個別に還元

# 3. 統計データ分析コンペティションの応募状況

## 受賞論文の傾向

これまで（2018～2022）の受賞論文中で用いられた分析手法について集計

### 高校生の部：

47本中37本が相関分析（他の分析手法との併用含む）  
回帰分析等高度な分析はそれほど多くない

### 大学生・一般の部：

53本中で用いられた分析手法は多岐に渡る  
相関分析、回帰分析、主成分分析、クラスター分析、  
パネルデータ分析、因子分析、ベイズモデリング…等

# 4. 受賞論文の概要等

## 受賞者及び受賞論文発表（10月18日）

### 「統計データ分析コンペティション2022」受賞者及び受賞論文 ～ 高校生の部 ～

受賞者	受賞論文(タイトル及び概要)
<b>【総務大臣賞】</b> <b>太佐 美結</b> (フェリス女学院高等学校)	<b>体力が基礎学力に与える影響について</b> 個別の体力テストの点数から抽出した体力因子が学力に与えている影響を分析し、家庭の経済力や学習環境なども含めた学力との関連性を明らかにすることで、小中学生の運動意欲の低下防止や、運動や学習を行う動機付けについて示した。
<b>【優秀賞】</b> <b>今泉 開</b> (慶應義塾高等学校)	<b>ボランティア活動の決定要因</b> 都道府県別のボランティア活動率を用い、重回帰分析や散布図を用いた分析を行うことで、ボランティア活動の決定要因が都道府県単位で存在し、その決定要因がそれぞれのボランティア活動に大きな影響を与えることを示した。
<b>【統計数理賞】</b> <b>林 蔚欣</b> (茨城県立並木中等教育学校)	<b>市区町村というミクロ的視点から投票率の実態を探る</b> 市区町村別の投票率を、得票数合計と人口から計算し、それを様々な変数で回帰することにより、離婚件数や財政力指数、完全失業率、民生費割合等がマイナス、教員数等がプラスの影響であることを見出した。
<b>【統計活用奨励賞】</b> <b>森下 達也</b> (愛知県立一宮高等学校)	<b>都市部と地方の教育格差の要因と課題</b> <b>～日本の教育現場において～</b> 都道府県別の進学者数、県民所得、教育費支出、教育関係事業所数等を用い、都道府県別の教育格差やその要因を分析するとともに、教育のデジタル化として、県別のデジタル教科書整備率や電子黒板整備率を比較し、都市部におけるデジタル整備が進んでいないことを示した。

※ 受賞論文については、後日以下のページで公表予定 <https://www.nstac.go.jp/statcompe/>

# 4. 受賞論文の概要等

## 「統計データ分析コンペティション2022」受賞者及び受賞論文 ～ 大学生・一般の部 ～

受賞者	受賞論文(タイトル及び概要)
<p><b>【総務大臣賞】</b></p> <p>大古 一聡、西川 直輝 (東京大学大学院情報理工学系研究科)</p>	<p>パンデミックは人流をどう変えたかー地域の特性別こー</p> <p>都道府県別の人流データを用い、移動距離やサービス業割合、旅館数等で重力モデルによる回帰分析を行い、新型コロナウイルス感染症による影響で長距離移動やサービス業・観光業の盛んな地域への移動が減ったことを示した。</p>
<p><b>【優秀賞】</b></p> <p>TENG YU ZHE (青森中央学院大学経営法学部)</p>	<p>社会保障政策と犯罪の関係</p> <p>ー都道府県パネルデータによる実証分析ー</p> <p>都道府県別データにより、犯罪件数を生活保護人員数割合、進学率、失業率等の変数で回帰し、社会保障政策が充実すれば犯罪が減ることを示した。</p>
<p><b>【統計数理賞】</b></p> <p>小関 敦生、黒須 咲菜、杉本 果穂、 守木 悠太郎、森田 花梨 (静岡大学情報学部)</p> <p>宇留賀 大誠 (静岡大学大学院総合科学技術研究科)</p>	<p>リサイクル活動に対する地域・政策要因の研究</p> <p>-主成分分析・階層的クラスタリングを用いた市町村別分析-</p> <p>市町村単位のデータを用い、主成分分析とクラスター分析を組み合わせた都市分類を行った上で、リサイクル率に関して重回帰分析を行い、産業規模の拡大が都市群によってリサイクル率に影響を及ぼすことを示した。</p>
<p><b>【統計活用奨励賞】</b></p> <p>原 明美 (東京理科大学理学部第二部)</p>	<p>生活系ごみ排出量と事業系ごみ排出量による回帰分析</p> <p>都道府県別パネルデータを用い、ごみ排出量と各種変数の関係を分析し、1人当たり県内総生産やごみリサイクル率が生活系ごみ排出量と事業系ごみ排出量にプラスの影響を与えることを見出した。</p>

- **統計データ分析コンペティション**  
<https://www.nstac.go.jp/use/literacy/statcompe>
- **SSDSE（教育用標準データセット）**  
<https://www.nstac.go.jp/use/literacy/ssdse/>
- **総務省統計局報道資料**  
<https://www.stat.go.jp/info/guide/public/houdou/pdf/ho220415.pdf>